STRATEGIC VOTING EQUILIBRIA UNDER THE SINGLE NONTRANSFERABLE VOTE

GARY W. COX *University of California, San Diego*

Previous investigations of strategic voting equilibria in mass electorates have looked only at single-member districts. I shall investigate such equilibria in multimember districts operating under the single nontransferable vote system. What appear to be the most natural equilibria conform to the M + 1 rule, according to which strategic voting in M-seat districts produces exactly M + 1 vote-getting candidates in equilibrium, any others having their support totally undercut. This result provides the beginnings of a formal underpinning for Reed's recent extension of Duverger's Law to the Japanese case. The model also generates specific and empirically testable hypotheses concerning the exceptions to the M + 1 rule that one ought to expect in equilibrium. I test these hypotheses with Japanese data. Finally, the model also reveals a type of strategic voting that is specific to multimember districts. I use Japanese data again to explore the empirical importance of this kind of strategic voting.

For as long as voting procedures have been used to decide important and controversial issues, there have been legislators and electors willing to vote strategically. Theoretical interest in strategic voting dates at least to Pliny the Younger (see Farquharson 1969) and probably earlier. I shall build on rather more recent and formal treatments of the strategy of voting: those framed in the decision-theoretic and game-theoretic traditions. Most of this work has appeared in the last 25 years and focuses on the behavior of legislators (e.g., Austen-Smith 1987; Banks 1985; Farquharson 1969; McKelvey and Niemi 1978; Miller 1980; Ordeshook and Schwartz 1987; Shepsle and Weingast 1984). I shall focus on the other, less well trodden area of research into strategic voting, that dealing with the behavior of voters in mass elections.

I deal in particular with elections satisfying the following three criteria: (1) each voter casts a single vote (for a candidate, not a party), (2) there are $M \geq 1$ seats to be filled in the district in which the election is held (the model focuses on one district at a time), and (3) plurality rule determines who wins the available seats (so the candidates with the M highest vote totals win election). Two electoral systems satisfy these criteria: the Anglo-American first-past-the-post system, in which $M = 1$, and the Japanese single nontransferable vote (SNTV) system, in which $M > 1$. The first of these systems is, of course, widely used; the second is currently employed at the national level only in Japan and Taiwan (and the Japanese are, as of this writing, in the midst of changing their electoral system).[1] Although structurally very similar, the difference in district magnitudes between the two systems leads to substantial practical differences in systemic performance and party strategy. Nonetheless, the two systems are alike in their strategic voting equilibria, as will be shown.[2]

Although my primary interest here is in the theoretical nature of strategic voting under SNTV (I shall consider the Anglo-American system to be SNTV with $M = 1$), some important and empirically testable implications of the theory arise along the way. These will be duly explored.

First, I shall review the previous formal literature on strategic voting (leaving the vast informal literature largely untouched). Then, I shall describe an *M*-seat *K*-candidate model of strategic voting under SNTV. Next, I shall show that in equilibrium, no votes are wasted at any margin, which entails both that identifiably weak candidates are deserted by their supporters (the feature upon which previous work has focused) and that identifiably strong candidates are deprived of their excess supporters. Finally, I shall explore the empirical usefulness of the model, using Japanese data as an illustration, and conclude.

PREVIOUS RESEARCH

The most widely recognized form of strategic voting occurs in single-vote plurality-rule elections held in single-member districts. As early as 1869, Henry Droop (an English advocate of proportional representation) recognized the basic logic: "As success depends upon obtaining a majority of the aggregate votes of all the electors, an election is usually reduced to a contest between the two most popular candidates. . . . Even if other candidates go to the poll, the electors usually find out that their votes will be thrown away, unless given in favour of one or the other parties between whom the election really lies" (quoted in Riker 1982, 756). Droop also saw an important systemic consequence of this district-by-district tendency toward strategic voting: the political nation was divided into two large catchall parties, rather than into many smaller and better-defined parties (see Riker 1982, 756–57). He thus anticipated what we now know as Duverger's Law—the proposition that electoral laws mandating the use of plurality rule in single-member districts tend to produce two-party systems (Duverger 1955; Riker 1982).

Formal mathematical study of strategic voting in the last 20 years has had two stages: an early decision-theoretic stage and a more recent game-theoretic stage. The decision-theoretic perspective on strategic voting (see McKelvey and Ordeshook 1972) is essentially the same as Droop's: some voter, whose favorite candidate has a poor chance of winning, notices a personal preference between the top two candidates; the voter then rationally decides to vote for the most preferred of these top two competitors, rather than for his or her overall favorite, because the latter vote has a much smaller chance of actually affecting the outcome than the former. The decision-theoretic approach adds to common sense not just greater precision about the assumptions implicit in Droop's reasoning (e.g., the probability not of victory but of ties and near ties matters directly) but also greater generality: the basic model has been extended to illuminate strategic behavior in multimember districts (Cox 1984), approval voting elections (Niemi 1984), and a variety of other electoral systems (Hoffman 1982).

Nonetheless, decision-theoretic analyses, both formal and informal, still deal essentially with a single voter in analytic isolation. The logical next step is to consider whether strategic voting by some voters makes such voting by others more or less likely. In particular, suppose a close third-place candidate in a single-member district begins to lose the support of his least committed followers (those who prefer him only slightly to one of the two frontrunners). This erosion of support will, if known (perhaps through polls), lead voters to reduce their estimates of the candidate's chances. But as the candidate's chances are seen to fall, some of his slightly more committed followers may abandon ship for one of the frontrunners. The process might, in theory, continue until the candidate was left with no support.

This line of thinking is game-theoretic. It essentially asks how much strategic voting there is in equilibrium. Should one expect that third-place candidates will always lose all of their support because of strategic decisions among their followers? Or are there general conditions under which this erosion of support is fairly limited or even negligible?

I addressed these questions earlier in the context of a model (similar to that in Ledyard 1984) in which three candidates compete for a single seat under the usual Anglo–American rules (Cox 1987). The key assumptions of the analysis were that all voters are instrumentally rational (i.e., they care whom they vote for only insofar as it affects the outcome of the election), that voters have incomplete information about each other's preferences over outcomes, and that all voters have "rational" expectations. I showed that in almost all equilibria some voters vote strategically and that the marginal impact of strategic voting was to decrease the effective number of parties (Laakso and Taagepera 1979).

Palfrey (1989), exploring essentially the same model, was able to characterize its equilibria in terms of candidate vote shares, showing that they fall into two classes: Duvergerian equilibria (in which the level of strategic voting is such that the support of all but two of the candidates is undercut completely) and non-Duvergerian equilibria (in which two or more candidates are so nearly tied for second that the voters cannot decide which one to discount, leaving more than two significant candidates in the field). Palfrey believed that the non-Duvergerian equilibria were razor's-edge phenomena and hence that he had provided an internally consistent explanation of Duverger's Law in terms of strategic voting alone. Myerson and Weber however, have shown that this belief is incorrect (1993, 106). Although the non-Duvergerian equilibria do seem unusual because of their requirement that two or more candidates be virtually tied for second, they are not unusual in the mathematical sense of being nongeneric. The intuition behind them is roughly as follows. Suppose two leftist candidates (e.g., Charles Goodell and Richard Ottinger) and one rightest (e.g., James Buckley) are competing for a single post (one of the U.S. Senate seats for New York in 1970). The rightist is ahead, the two leftists trailing but close to one another. Under these conditions, leftist voters will have a hard time coordinating on one of the leftist candidates, and a non-Duvergerian result can (and did) ensue. This failure to coordinate, moreover, does not require nongeneric conditions on the distribution of voter types, such as that the two leftists have exactly the same number of voters ranking them first. If more voters rank leftist A first than rank leftist B first, for example, then the two can still end in a dead heat (or sufficiently close) if it is expected that B has a larger chance of being tied for first with the rightist than does A. Such beliefs appear unusual in that they require the "objectively weaker" candidate B to have a better chance of being tied for first than the "objectively stronger" candidate A, but they are not ruled out by the Bayesian equilibrium concept. At present, therefore, Duverger's Law cannot be derived exclusively from analyses of strategic voting equilibria.[3]

Myerson and Weber advance a model of voting equilibria applicable in a wide range of single-winner electoral systems—not just ordinary plurality rule but also approval voting, Borda's method of points, and many other systems as well. The main difference between their approach and Palfrey's is that they assume as an axiom something Palfrey derives endogenously, roughly, that candidates generally expected to place third or lower in the poll are much less likely to be tied for first than candidates generally expected to place first or second in the poll.

I shall try to extend the study of strategic voting to cover multimember districts. I adapt Palfrey's model to the multimember context, holding all other institutional features (one nontransferable vote per voter, seats allocated by plurality rule) constant. At the same time, some features of the Myerson and Weber (1993) and Hoffman (1982) approach are incorporated as well.

A *K*-CANDIDATE *M*-SEAT MODEL

There are K candidates, indexed by $\boldsymbol{K} = \{1, \ldots, K\}$, competing for M seats. The M candidates placing highest in the poll win the available seats. If two or more candidates are tied for Mth in the poll, then the tie is broken equiprobably. After any ties have been broken, the *outcome* of the election is a set of M elected candidates. The set of all possible outcomes, denoted Ω, accordingly has $\binom{K}{M}$ members. Candidates are nonplayers in the model; they take no actions and are simply "entities with fixed characteristics about which voters might care."

There are n voters, each possessing a (strict) preference ranking over the possible outcomes of the election. The analysis is simplified, without altering the basic point, if we assume that voters' preferences over outcomes are additively related to their preferences over candidates, that is, each voter i is able to assign Von Neumann–Morgenstern utilities to each candidate—$u_{i1}, u_{i2}, \ldots, u_{iK}$—such that outcome $\alpha \in \Omega$ is preferred to outcome $\beta \in \Omega$ if and only if $\Sigma_{j\in\alpha} u_{ij} > \Sigma_{j\in\beta} u_{ij}$, with indifference obtaining if and only if the summed utilities are equal. Voter utilities can be rescaled in the standard fashion so that the most preferred candidate yields a utility of 1, the least preferred candidate, a utility of 0. After this rescaling, voter i's preferences (or voter i's *type*) can be described by the vector $u_i = (u_{i1}, \ldots, u_{iK})$, an element in the set $U = \{(u_1, \ldots, u_K): \max\{u_j\} = 1 \ \& \ \min\{u_j\} = 0 \ \& \ u_j = u_k \text{ only if } j = k\}$.

The relative frequencies of the different possible types of voter preference are given by a cumulative distribution function F defined over U. I assume that all possible preferences are represented in the electorate, possibly with vanishingly small probability (formally, the support set of F is U). I also assume that the probability two randomly sampled voters have identical preferences is negligible (formally, F has no mass points).[4]

Each voter chooses a vote in order to maximize expected utility, something that depends not just on the voter's preferences over candidates but also on his or her *expectations* about how well each candidate is likely to do. These expectations are formalized here as a vector $\pi_i = (\pi_{i1}, \ldots, \pi_{iK})$, where π_{ij} equals i's (subjective) probability that a randomly selected voter will vote for candidate $j \in \boldsymbol{K}$. Equivalently, π_{ij} is the expected proportion of the electorate who will vote for j, according to i. Given preferences (u_i), expectations (π_i), and knowledge of the number of voters (n), voter i faces a standard decision problem, whose details are given in Appendix A. The solution to i's problem (i.e., the set of votes that maximize expected utility, given u_i, π_i, and n) is denoted $V(u_i; \pi_i, n) \subseteq \boldsymbol{K}$.

The model is completed with two further assumptions whose joint effect is to restrict the nature and consistency of voter expectations. First, I assume that the distribution of voter preferences, F, is common knowledge. Second, I assume that voters' expectations are publically generated (e.g., by polls and newspaper analysis of the candidates' chances), so that diversity of expectation among the electorate is minimized. In the discussion that follows, I take this notion to the logical extreme and assume that every voter has the same expectations: $\pi_i = \pi$ for all i. This assumption can be replaced with a less stringent one, however, namely, that the expected *order* of finish of the candidates is the same for all voters, in the sense that they all agree on which candidates are *trailing* (expected vote shares that put them strictly below (M + 1)th place), which are *leading* (expected vote shares that put them strictly above Mth place), and which are *marginal* (neither trailing nor leading).[5]

Given these two postulates, the maintained assumption of voter rationality implies a certain consistency between F and π in equilibrium, for, not all expectations π are "rational" in light of the voter's knowledge of the distribution F of voter preferences. Suppose (to take a four-candidate example) that some voter thought π equaled (.25, .25, .25, .25), so that a randomly selected voter was equally likely to vote for any of the candidates. This expectation is clearly not consistent with a distribution of voter preferences in which the proportion of voters ranking candidate 1 last exceeds .75, since voting for 1 is a dominated strategy for such voters.

More generally, if π is a publically generated expectation common to all voters, and the number of voters is common knowledge, then each voter can simply compute the optimal votes $V(u; \pi, n)$ of all other voter types u. Then, knowing the distribution of voter types, the voter can calculate the proportion of the electorate who, given π and n, will vote for each candidate.

To see this, let $H_j(\pi, n) = \{u \in U: j \in V(u; \pi, n)\}$ be the set of all voter types who will vote for j, given π and n. Then the probability that a randomly sampled voter will vote for j is simply $\int_{H_j(\pi, n)} dF$.[6] If $\int_{H_j(\pi, n)} dF$ is not equal to π_j, then the original expectations are not tenable: to continue to believe them entails believing that some other voters will not vote rationally.

This is the gist of the argument for imposing the following "rational expectations" condition on voter beliefs.

Rational Expectations Condition. *The expectations π are rational with respect to the distribution F if, for all j,* $\pi_j = \int_{H_j(\pi, n)} dF$.

The equilibrium conditions for the model are then two. First, every voter votes so as to maximize expected utility, given expectations π (and n); that is, voters of type u vote for a candidate in $V(u; \pi, n)$. Second, the expectations π satisfy the rational expectations condition. These equilibrium conditions are identical to those imposed in a symmetric Bayesian Nash equilibrium, the only differences being the interpretation of the model's elements.

VOTING EQUILIBRIA & WASTED VOTES

What are the equilibria of the model just specified? Letting $N(\pi) = |\{j \in \boldsymbol{K}: \pi_j > 0\}|$ be the number of candidates with positive expected vote shares, there are three main cases to consider: $N(\pi) < M$, $N(\pi) = M$, and $N(\pi) > M$.

Equilibria with N(π) < M. Equilibria with fewer vote-getting candidates than there are seats to be had do not exist. Suppose, to the contrary, that $N(\pi) < M$ for some expectations π satisfying the rational expectations condition. In this case, perversely enough, *no* voter would actually vote for any candidate j such that $\pi_j > 0$ (i.e., $\pi_j > 0 \rightarrow j \in V(u; \pi, n)$ for no u). This follows because, given π, n, and the usual Nash assumption, each voter believes that all candidates j with $\pi_j > 0$ are certain to be elected. Thus there is no incentive to vote for them. On the other hand, voting for the most preferred of the candidates j with $\pi_j = 0$ is, under the same assumptions, certain to elect that candidate (by breaking a tie among K–$N(\pi)$ candidates, all with zero votes, in favor of the selected candidate).

Equilibria with N(π) = M. Such equilibria do not exist either. Consider a four-candidate, two-seat example to illustrate the point. If $\pi = (0, 0, .5, .5)$ is an equilibrium, then it must be that a voter preferring candidate 1 to 2, 2 to 3, and 3 to 4 would vote for 4 or for 3, given π. But voting for 4 is a dominated strategy for this voter. And given π, voting for 3 is not optimal, for it is possible that 3 will receive less than two votes. In these cases, a vote for 1 is superior to a vote for 3, since it either breaks a three-way tie for second in 1's favor (if 3 receives no votes) or puts 1 into a two-way tie with 3 for second (if 3 receives one vote), both desirable results. In all other cases (i.e., 3 has more than one vote) voting for 1 and voting for 3 yield identical outcomes. Thus, no voter of the specified type will vote for either 3 or 4, and given the assumption that there are some voters of each type, the assumption that π satisfies the rational expectations condition is contradicted.

Equilibria with N(π) > M. This is the most interesting case, and the one that takes the most work to nail down. Relabel the candidates, if necessary, so that $\pi_j \geq \pi_{j+1}$ for all $j \in \mathbf{K}\backslash\{K\}$. Note that with this relabeling, the condition $N(\pi) > M$ implies that $\pi_{M+1} > 0$. Given a distribution F of voter types, I shall say that the expectations π are a limit of rational expectations if and only if, for every $\varepsilon > 0$, there exists an integer N and a sequence $\{\pi^n\}$ of rational expectations such that $n > N$ implies $|\pi_i^n - \pi_i| \leq \varepsilon$; that is, π is a limit of rational expectations if and only if arbitrarily large electorates can have rational expectations that are arbitrarily close to π. The main result is presented in the following theorem and its corollary.[7]

THEOREM 1. *Suppose that either (1)* $0 < \pi_j < \pi_{M+1}$ *for some* $j > M + 1$ *or (2)* $\pi_M < \pi_1$. *Then* π *is not a limit of rational expectations.*

For the proof, see Appendix B. The basic logic of the proof is this: if $0 < \pi_j < \pi_{M+1}$, then candidate j is virtually sure to lose for sufficiently large n, and voting for the most palatable of the candidates most likely to be tied for Mth yields a higher expected utility than voting for j; if $\pi_M < \pi_1$, then candidate 1 is almost certain to win for sufficiently large n, and, again, the instrumental voter's attention turns to those candidates with "asymptotically large" conditional tie-probabilities.

A direct corollary of Theorem 1 is the following.

COROLLARY 1. *If* π *is a limit of rational expectations, then (1)* $\pi_1 = \pi_2 = \ldots = \pi_M$ *and (2)* $\pi_j \in \{0, \pi_{M+1}\}$ *for all* $j > M + 1$.

For $M = 1$, this corollary simply extends Palfrey's earlier three-candidate result to cover an arbitrary number K of candidates. For $M > 1$, the corollary carries two messages about wasted votes, rather than just one. Like Palfrey's previous work, it shows that strategic voting works against trailing candidates (those who fall "too far" behind in the polls). In addition, the corollary shows that strategic voting relieves leading candidates (those who are "too far" ahead in the polls) of their excess votes, reducing them to equality with all other likely winners. Thus votes are wasted neither on weak nor on strong candidates.

The nature of the two processes (winnowing out the weak, equalizing the strong) is slightly different. Voters who find themselves supporting a trailing candidate are more likely to desert that candidate as other voters desert him. There is thus a certain momentum to strategic voting as it affects weak candidates. Voters who find themselves supporting a leading candidate, in contrast, are less likely to abandon that candidate upon hearing of other defections. To illustrate this, consider a contest between three candidates (A, B, and C) in a two-seat district. Suppose A is ranked first by 40% of the electorate and is expected to get all of their votes, while B and C are each expected to get 30% of the vote. In this case, A voters who prefer C to B are tempted to switch their support to C, while A voters who prefer B to C are tempted to switch their support to B. If those who prefer C (B) succeed in bringing A and C (B) into a dead heat at 35% (by abandoning A), then similar action by those who prefer B (C) will be forestalled. If both sides (the A supporters who prefer B to C and those with opposite preferences) abandon A, then the worst outcome for both results. Thus, there is a Chicken aspect to the game: as soon as the existence of excess votes becomes common knowledge, both sides begin "driving" toward abandonment. The question is, Which side will swerve first? The present model, it should be noted, does not capture this Chicken aspect at all well formally, since the equilibrium concept simply assumes away all coordination difficulties. This may be one important reason that the model's predictions regarding the elimination of votes wasted on strong candidates fare so poorly in the empirical analysis.

As in the single-seat case, the corollary divides multiseat equilibria into two classes: (1) Duvergerian equilibria, with $M + 1$ vote-getting candidates, and

(2) non-Duvergerian equilibria, with more than $M + 1$ vote-getting candidates. The Duvergerian equilibria entail a close M-way race for the M available seats, with a single runner-up, all other candidates being reduced to near-zero support. The non-Duvergerian equilibria look the same, except that there are two or more runners-up, whose nearly identical expected vote totals prevent any being winnowed out from the field of viable candidates.

STRATEGIC VOTING AS AN EXPLANATION OF REAL-WORLD DATA

I shall consider the empirical usefulness of the results just sketched. There is no question that instrumentally rational agents of the type stipulated, with rational expectations, will behave in a very precise fashion. But of course it is possible to doubt that real people are entirely instrumentally motivated or that they have rational expectations. And it is a matter of simple observation that the top M candidates in SNTV elections do not have virtually identical vote shares. Taken as is, then, the model's predictions meet with immediate empirical disconfirmation. A reasonable conclusion is that one or more of the model's assumptions are poorly approximated in the real world.

Overly precise predictions are typical of highly abstract models and a typical (often unstated) assumption of theoreticians is that the model's predictions could fairly easily be made more reasonable, without changing their qualitative nature, by adding a bit of "noise" or "friction" to the model. Since the empirical testing of rational choice models has been strongly criticized of late (Green and Shapiro 1993), I will explore what some of the noise to be added might be. Even if adding noise (e.g., noninstrumental voters or voters whose expectations are inconsistent) can in principle produce predictions not obviously false, there is still interest in two questions: (1) Do real-world data conform sufficiently closely to the model's predicted equilibria that one might believe that a model essentially similar to this one (just adding noise) might tally with real-world patterns? (2) Even if the real world conforms to sylized versions of the model's equilibria, are there other explanations that predict the same patterns? I shall examine each of these topics—noise, empirical patterns, and alternative explanations—in turn.

Noise

There are many frictions that would, if introduced into the model, soften its predictions. Here, I shall discuss only two: the introduction of noninstrumental and of "ignorant" voters. Introducing voters who are not purely instrumental but also derive utility from the particular action that they take (abstaining, voting for some candidate) can have various implications. If one assumes that voters bear some cost from voting but that conditional on voting, they do not care whom they vote for except insofar as it affects the outcome, then not much changes. One would have to assume that a fixed proportion of the electorate had nonpositive costs of voting (or something to this effect) in order to generate growing turnout, but after that, the model's results would go through unchanged.

On the other hand, if one assumes that voters care about whom they vote for as well as who wins, then much may change. For such *mixed-motivation* voters, action-contingent utilities would dwarf outcome-contingent utilities as the electorate grows because outcome-contingent utilities are discounted by the probability of a single vote affecting the outcome, whereas action-contingent utilities are not. Thus, a mixed-motivation voter who votes at all votes for the candidates that he or she likes to vote for, regardless of whether this is the candidate that "strategic voting" would dictate supporting. If the entire electorate had mixed motivations, then expectations about how well the candidates were doing would be of negligible importance to voter decisions, and *any* pattern of expectations could be consistent with voter rationality in equilibrium.

If only some voters have mixed motivations, then the result is more interesting. Let Q be the proportion of the electorate that is purely instrumental or outcome-oriented and let $1 - Q$ be the proportion with mixed motivations. In the limit, the latter group will be purely noninstrumental or action-oriented, as noted. Let F be the distribution of preferences in the instrumentally motivated electorate and let P_j be the proportion of the noninstrumentally motivated electorate that prefers voting for candidate j. Then the rational expectations condition becomes $\pi_j = Q \int_{H_j(\pi, n)} dF + (1 - Q)P_j$ for all j. If candidate j is trailing, then j loses all *instrumental* support but retains support among the mixed electorate. Thus, if Q is close to 1, the result does not change much: in equilibrium, strategic voting will reduce weak candidates—if not to zero support, then down to some fairly small minimum (i.e., $(1 - Q)P_j$) of noninstrumental support and will reduce strong candidates down to equality with all other likely winners.

Introducing "ignorant" voters relaxes a different premise of the model, namely, that the identity of trailing, marginal, and leading candidates is common knowledge. The publicness of this knowledge keeps all instrumental voters on the same page of the playbook: they *all* desert the (publicly identified) trailing and leading candidates in order to focus on the (publicly identified) marginal candidates.

One might argue for the reasonableness of the common-knowledge assumption as regards trailing candidates by noting the self-fulfilling character of voter expectations. If every voter *believes* that candidate j is out of the running, then j will in fact *be* out of the running. Moreover, if some voters, who previously intended to vote for j, come to believe that j is behind, they will desert j, thereby making it more likely that j is behind.

The arguments just given do not really justify *assuming* that the identity of trailing candidates is common knowledge, however. They only justify a belief that in equilibrium, the identity of trailing candidates will probably be common knowledge. To simply assume the common-knowledge condition is similar to assuming that the players in a two-person Battle of the Sexes will coordinate on one of the two pure-strategy Nash equilibria. As regards leading candidates, voter expectations are not even self-fulfilling. If each voter believes that candidate j is leading (hence virtually sure to win), then each will abandon j, who will in fact finish out of the running.

If who trails and who leads is not common knowledge, then an extra degree of freedom is opened up in the model. In the extreme, the analyst can stipulate (possibly inconsistent) expectations for each voter. This degree of analytical latitude would be enough to make any pattern of aggregate vote returns consistent with some equilibrium of the model. On the other hand, placing limits on the extent to which voters' expectations differed would begin to restore some "bite" to the model's predictions.

These observations motivate asking how voters learn about the candidates' expected vote shares. In the real world, the forces generating common knowledge of candidate chances are polls, news analyses, candidate statements, and other bits of essentially free information. It has to be free information because rationally ignorant voters will not exert any effort to determine who is ahead for the same reason that they will not research candidate positions carefully (Downs 1957). Thus, the extent to which the real world approximates the model's strictures should depend on the availability and clarity of free information regarding the relative standing of the candidates. If voters are exposed to lots of free information (e.g., frequently published polls) that reveals some candidates to be clearly trailing the others and if this information seeps out to a large proportion of the instrumental electorate, then one expects that trailing candidates will be left with not much more than their noninstrumental support. If voters have no information regarding candidate chances (and diffuse priors), then sincere voting is consistent with expected utility maximization, and one does not expect objectively trailing candidates (those who have fewer voters ranking them first) to lose their instrumental support. If (to take a third example) voters have conflicting information regarding candidate chances, then strategic voting by some voters may "cancel out" strategic voting by others, leaving little or no observable impact on the aggregate distribution of votes.

From this perspective, the tendency of candidates trailing in multicandidate races to dispute the accuracy of the polls that show them trailing, to claim to have different results in proprietary polling, and to urge voters to ignore the polls is understandable. All these actions make good sense from the point of view of preventing their last-place status from becoming common knowledge.[8]

Empirical Patterns

The model predicts that strategic voting will produce two main patterns: (1) trailing candidates will be deserted by all their supporters, and (2) leading candidates will be deserted by their excess supporters. In examining how these two predictions fare, I shall look at postwar Japanese electoral data from the period 1958–90. Evidence is fairly good that Japanese voters do strategically desert weak candidates. Steven Reed (1991), in an extensive examination of the postwar Japanese experience with SNTV, finds a clear and increasing tendency for there to be only $M + 1$ viable candidates in each district. This pattern, of course, fits the Duvergerian equilibria identified by the model.

What of the non-Duvergerian equilibria? These equilibria all entail that the first and second losers receive nearly the same number of votes. Thus a theoretically interesting statistic is the *second*-to-*first* loser's vote total ratio (SF ratio). Under Duvergerian equilibria, the SF ratio will be near 0. Under non-Duvergerian equilibria, the SF ratio will be near 1. Thus, if one were to compute the ratio for a number of districts and plot the resulting distribution, one should find a spike at 0 and a spike at 1.

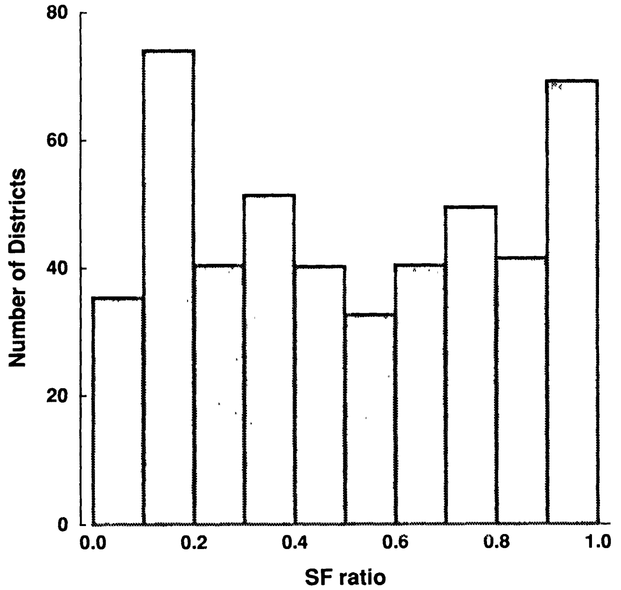
Allowing for some frictions in the model (e.g., some noninstrumental voters, some disagreement about which candidates are trailing or leading, and which are marginal), the prediction is softened. The SF ratio should either be close to 1 (when second losers are so close in the polls to first losers that they do not lose their support due to strategic voting) or close to 0 (when second losers are sufficiently far behind first losers that strategic voting kicks in and they are reduced to their noninstrumental support level, which I assume to be close to 0 for most candidates). The SF distribution, in other words, should be bimodal.

I have tested this bimodality hypothesis empirically in the case of Japan, using district-level electoral returns over the period 1958–90.[9] The procedure, in the case of three-seat districts, was as follows. First, I computed the ratio of the vote total of the second loser (fifth-place candidate) to the vote total of the first loser (fourth-place candidate) for all districts with at least five candidates. Then, I produced a histogram to summarize the distribution of the resulting SF ratios (Figure 1). Results for four- and five-seat districts (the other frequently occurring types of district in Japan) are given in Figures 2 and 3.

As can be seen, the SF distribution is bimodal in each case. Values near .5 are rare, relative to those near 1 or 0; that is, it is much more common to have either a close or a distant second loser than an "in-between" second loser. Moreover, the closer the first loser to the last winner, the more likely a few more votes might change the outcome, the further

FIGURE 1

Testing the Bimodality Hypothesis in 3-Seat Districts: Japanese Lower House Elections, 1958–90



Note: The SF ratio in any given district is found by dividing the second loser's vote total by the first loser's vote total.

from .5 is the SF ratio (i.e., the stronger is the tendency for the ratio to be either near 1 or near 0).

Are the distributions displayed in Figures 1–3 significantly bimodal? One can reject the null hypothesis that the distribution is unimodal in the first two cases, using dip or depth tests (see Hartigan and Hartigan 1985). In the third case, the probability of observing the degree of bimodality visible in Figure 3 is a bit below .2, under the null hypothesis that the distribution is really unimodal. Thus there is some reason to doubt that the distribution is unimodal, but the evidence is not conclusive at conventional levels of statistical significance. Using a kernel density-based test proposed by Silverman (1981), one can pit the null hypothesis of *bi*modality against the compound alternative of more than two modes. Doing so, one finds p-values of .22, .26, and .98 for the first, second, and third figures respectively. Thus one cannot reject the null of bimodality (in favor of multimodality) at conventional levels of significance. All told, the evidence is as it appears to be to the naked eye: the first two figures really are bimodal, and the third is harder to call but has some tendency toward bimodality.

Substantively, does all this mean that one can reject the null hypothesis of sincere voting? It all depends on what one believes the distribution of voters' true preferences over candidates is. If one believes that second losers' true support tends to be either almost equal to, or much less than, first losers' true support, then the evidence presented here is consistent with sincere voting. I can see no reason for such an expectation regarding the distribution of preferences, however, and accordingly take the statistical tendency toward bimodality to be evidence of strategic voting.

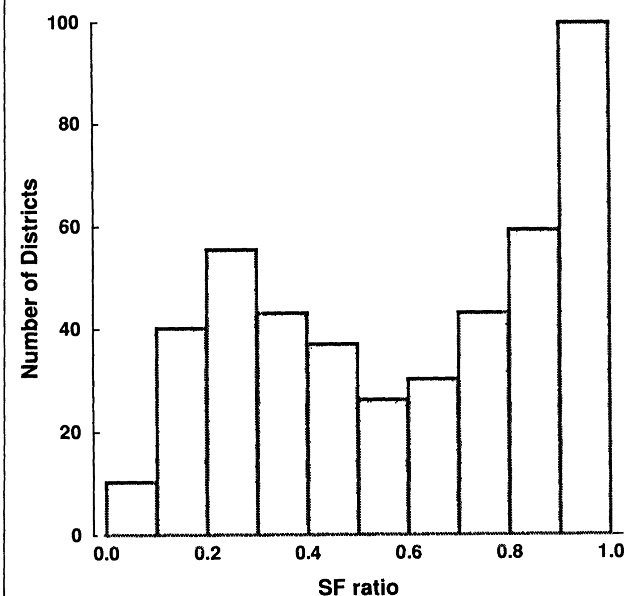
A second-order pattern that appears in the data is that the height of the mode near 0 declines with district magnitude. My interpretation of this is that the quality of voter information regarding candidate chances declines with district magnitude. In particular, it is harder to be sure who is trailing in a more crowded field in which very small vote percentages can win a seat. Whether this notion will hold up to scrutiny is a topic for further investigation.

Evidence that Japanese voters strategically desert leading candidates (those with more support than they need) is much less compelling than that they desert weak candidates. There is rarely anything like a dead heat among the top M candidates in a Japanese district, nor is there any movement toward such an ideal over time. Nonetheless, there is some evidence consistent with the model: (1) there is an over-time trend within the dominant Japanese party, the Liberal Democrats, toward fewer seats being lost on account of votes "wasted" on strong candidates (Cox and Niou 1994); (2) there is a statistically discernible tendency for fewer votes to be "wasted" on leading candidates when the margin of victory of the last winning candidate is narrower (indicating that votes switched from leading to marginal candidates are more likely to affect the outcome).

This last point is documented via an ordinary least squares regression, the results of which are reported in Table 1. The dependent variable, EXCESS, measures how far a particular district's race departs from the

FIGURE 2

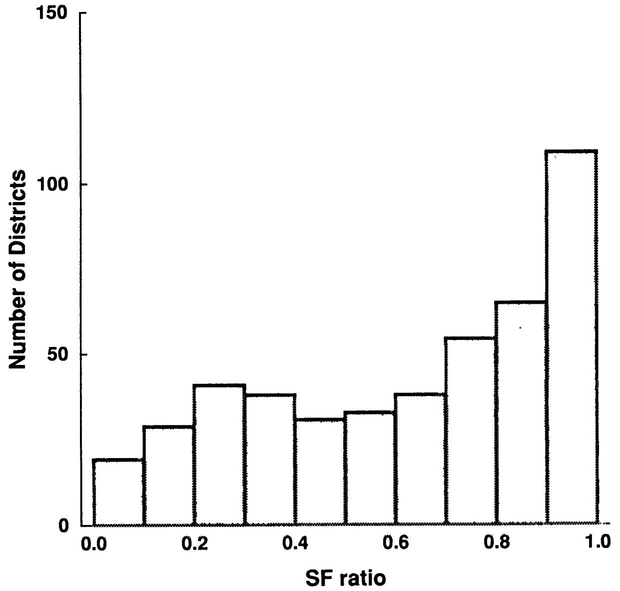
Testing the Bimodality Hypothesis in 4-Seat Districts: Japanese Lower House Elections, 1958–90



Note: The SF ratio in any given district is found by dividing the second loser's vote total by the first loser's vote total.

FIGURE 3

Testing the Bimodality Hypothesis in 5-Seat Districts: Japanese Lower House Elections, 1958–90



Note: The SF ratio in any given district is found by dividing the second loser's vote total by the first loser's vote total.

theoretical prediction that the vote percentages of the top M candidates will be the same. It is computed by calculating, for each of the top $M - 1$ finishers, the excess vote percentage they garner over that obtained by the Mth-place finisher and then adding the resulting $M - 1$ figures. For example, in a three-seat district in which the top three candidates get 30%, 25%, and 20%, EXCESS = (30 − 20) + (25 − 30) = 15. The main independent variable in the analysis is MARGIN, equal to the vote percentage of the last winning candidate less that of the first losing candidate. The theoretical expectation is that smaller margins of victory will lead to fewer excess votes, since smaller margins indicate a better chance that votes currently wasted on leading candidates might affect the race between the marginal candidates, if transferred there.[10] Because five-seat districts have four candidates above the Mth place while four-seat districts have three and three-seat districts only two, one expects more excess votes in larger-magnitude districts. It thus makes sense to control for district magnitude, which is done via two dummy variables.

Examining the results diplayed in Table 1, one finds a positive and statistically significant effect: a 1% reduction in the margin of victory of the last winner produces, on average, nearly a .5% reduction in votes wasted on leading candidates. There is thus a mild tendency for closeness in the race for the last seat in a district to reduce the number of votes wasted on strong candidates, even if there are still quite a few such wasted votes cast.

Alternative Explanations

Although the model of strategic voting generates empirically testable predictions, some of which are new (in the sense that they have not been noticed in the previous literature), there are also some obvious alternative explanations that might explain the pattern of evidence just uncovered. The problem is that *any* class of agents who care about the outcome of the election (not just voters but also activists, contributors, and candidates) will tend to allocate whatever resources they control (labor, money, etc.) to marginal candidates, where they are more likely to affect the outcome, rather than to trailing or leading candidates, where they are less likely to affect the outcome. Moreover, allocation or reallocation of resources to marginal candidates should produce the clearest aggregate results (trailing candidates deprived of all instrumental support, leading candidates of excess instrumental support) when who is marginal and who is nonmarginal is widely agreed and the margin of victory is small. Thus, the empirical evidence adduced is far from proving that a significant proportion of the electorate is instrumentally motivated and in reasonable agreement on the relative standing of the candidates. It may be that contributors give only (or mostly) to marginal candidates, that trailing candidates try to sell their endorsement to marginals, or that leading candidates sell some of their excess support to marginals.

The elite-level hypotheses have some distinct advantages. First, it is more plausible that elite actors, having larger stakes in the outcome, will pay attention to how close the race is and respond by diverting resources to marginal candidates. Put negatively, it is unlikely that ordinary voters will pay any attention at all, since their single votes have an infinitesimal chance of affecting the outcome. If it is at all costly to

TABLE 1

Closeness Depresses Excess Votes: Japanese Lower House Elections, 1958–90

INDEPENDENT VARIABLE	DEPENDENT VARIABLE: EXCESS — COEFFICIENT ESTIMATE	STANDARD ERROR
Constant	7.27	.35
Margin	.44	.04
Dummy for 4-seat districts	2.77	.42
Dummy for 5-seat districts	5.41	.43

Note: Adjusted R^2 = .12; N = 1,483. The variables are defined as follows. The unit of observation is an election in a particular district. Let V_j be the vote percentage garnered by the candidate placing jth in the poll in that district and let M be the district magnitude. Then (1) Excess = $\Sigma_{j=1}^{M-1} (V_j - V_M)$, (2) Margin = $V_M - V_{M+1}$, (3) Dummy for 4-seat districts = 1 if $M = 4$ (0 otherwise), and (4) Dummy for 5-seat districts = 1 if $M = 5$ (0 otherwise). As an example of the computation of Excess, consider a three-seat district in which the top three candidates receive 30%, 25%, and 20% of the vote. In this district, Excess = (30 − 20) + (25 − 20) = 15. The equation is estimated by ordinary least squares.

find out who is marginal or to calculate expected utilities, rational voters should avoid these costs, since bearing them has virtually no impact on the outcome (Meehl 1977; Riker 1982).

Second, the elite models look as if they can better accommodate the rather substantial amount of votes wasted on leading candidates. Contributors who seek "access" or specific postelection favors will stick with leading but not with trailing candidates. Leading candidates have something to lose if they sell off "too much" excess support, while trailing candidates, for all intents and purposes, do not.

Nonetheless, despite the apparent advantages of elite-based models, it is not clear that one can reject the voter-based model. The informational and cognitive costs of strategic voting are modest and may be borne entirely as byproducts of everyday activities, such as reading the newspaper, watching television, or attending college courses in politics. Information on the relative standings of candidates is sometimes published in polls; it does not take a rocket scientist to understand traditional wasted-vote arguments, and these arguments are sometimes hard to avoid (being urged by concerned elites). All this suggests that voters in the real world may strategically desert weak candidates for essentially the reasons stylized in the model. It is true that the whole process is mediated by elites: they point out that the race is close and that votes on weak candidates are wasted. But the voters do the rest: they buy the argument and act accordingly. Empirically, I think that there is overwhelming evidence that voters have, in a variety of historical and electoral contexts, voted strategically in this sense (see Cox 1991). The question of the relative importance of strategic reallocation of votes in the mass electorate as opposed to strategic reallocation of other resources in the elite strata remains open, however. And the evidence for strategic desertion of leading candidates is another story entirely. Elites in Japan do not seem to point out that votes for a sure winner are wasted, and there is no direct evidence at present that voters act as if they knew.

CONCLUSION

I have investigated strategic voting equilibria in multimember districts operating under SNTV, building on the previous work of Cox (1987), Palfrey (1989), and Myerson and Weber (1993). "Strategic voting" in single-member districts refers to a voter deserting a more preferred candidate with a poor chance of winning for a less preferred candidate with a better chance at winning. In multimember districts, voters who care only about the outcome of the election will strategically desert both candidates who are "too weak" and candidates who are "too strong." Such outcome-oriented voters desert weak candidates in multimember districts for the same reason as in single-member districts. They desert strong candidates when those candidates have one of the M seats sewn up but there are other seats still up for grabs; for then the voter's vote has a much greater chance of affecting the outcome if cast for one of the "marginal" candidates—those on the edge between winning and losing. All told, instrumentally motivated voters under SNTV waste their votes neither on weak (submarginal) nor on strong (supermarginal) candidates.

Equilibria of a "frictionless" model of strategic voting under SNTV—in which all voters are instrumentally motivated and all have rational expectations—are such that all wasted votes are wrung out of the system. As far as wasting votes on supermarginal candidates goes, this means that all winning candidates must be in a virtual dead heat in equilibrium, so that none will be perceived as having any "excess" votes that might profitably be transferred elsewhere. This prediction fares poorly in the Japanese data, although there is a tendency for fewer votes to be wasted on supermarginal candidates as the gap between last winner and first loser narrows.

Wringing out all votes wasted on submarginal candidates produces either Duvergerian outcomes (in which strategic voting erodes the support of all but one serious challenger, so that there are $M + 1$ viable candidates all told) or non-Duvergerian outcomes (in which two or more serious challengers are so closely matched that none becomes the unique victim of strategic desertion and more than $M + 1$ viable candidates survive). Empirically, outcomes with $M + 1$ viable candidates occur rather frequently. Thus, Duverger (1955) saw a tendency toward two-partyism in electoral systems employing single-member districts, while Reed (1991) saw a tendency for there to be four viable candidates in three-member districts, five viable candidates in four-member districts, and six viable candidates in five-member districts.

Both Duverger and Reed, of course, appealed to strategic voting (among other things) in explaining the observed tendency toward $M + 1$ viable candidacies (the $M + 1$ rule). The present model's utility is in illuminating some of the logical prerequisites and consequences of explaining the $M + 1$ rule in terms of strategic voting. As to prerequisites, the model shows that the degree to which strategic voting will winnow out weak candidates depends on how many instrumentally motivated voters there are and on how consistent their expectations about the relative standings of the candidates are. The empirical approximation of both these conditions plausibly depends on elite action and propaganda. American third-party movements (Ross Perot included) frequently emphasize *future* election outcomes: "We may have no real chance this time," they say, "but vote for us anyway, send a message, and help restructure American politics." The established party most hurt by the third party's appeals, in turn, is apt to emphasize the electoral here and now—the instrumental motivations highlighted in the present model. Similarly, elite actions determine how consistent voter beliefs are regarding who is winning and who is losing. If clear information about candidate chances is provided to voters, one can expect substantial strategic voting and a consequent reduction in the number of viable candidacies. If little (or conflicting) information is provided to voters, then greater amounts of sincere voting (or cross-cutting strategic voting) can be expected, and the tendency toward $M + 1$ viable candidates will be weaker.

As to the logical consequences of explaining the $M + 1$ rule in terms of strategic voting, the model provides specific and empirically testable predictions about what kind of exceptions to the $M + 1$ rule one should expect. Neither Duverger nor Reed do this. Both readily admit the possibility of exceptions to their generalizations: Duverger's Law and Reed's extension of it are both stated as "tendencies." But neither says much about the nature of the exceptions. Here, the theoretically allowable exceptions to the $M + 1$ rule have been characterized as various kinds of near ties, and a general implication of these exceptions (embodied in the bimodality hypothesis) has been tested with Japanese data.

There are of course other possible avenues to explore in explaining the $M + 1$ rule. Both Duverger and Reed appropriately suggest that elites may get into the act. Meehl (1977) and Riker (1982) argue that voters have too small a stake in elections to motivate strategic voting and emphasize elite actors even more strongly. Here I have noted that strategic reallocation of resources by outcome-oriented elite actors (activists, contributors, candidates) should produce many of the same aggregate patterns as identified in the voters-only model. My personal bias is strongly toward the elite-level hypotheses, as it is in the study of turnout (Cox and Munger 1989). I think strategic voting survives, both in theory and in practice, because one of the things outcome-oriented elites can do in close races to reallocate resources from nonmarginal to marginal candidates is to flood the mass media with wasted-vote arguments (including therein both the relevant evidence on candidate standings and the basic logic motivating a strategic vote). Finally, it should be noted that the basic model constructed here can be adapted to other electoral systems operating in multimember districts—in particular, to various forms of proportional representation. I intend to demonstrate this more fully in future work.

APPENDIX A: THE VOTER'S DECISION PROBLEM

I consider how a voter motivated solely by a desire to affect the outcome of the election decides whom to vote for, given that the voter votes. There are three parameters in the voter's decision (subscript is are suppressed): (1) the voter's preferences over the candidates, given by $u = (u_1, \ldots, u_K) \in U$; (2) the voter's expectations about how well each candidate will do at the polls, given by $\pi = (\pi_1, \ldots, \pi_K)$; and (3) the number of voters, n. I shall denote by $V(u; \pi, n) \subseteq \boldsymbol{K}$ the optimal vote(s) of a voter of type u facing an electorate described by π and n. I shall show that the parameters identified are indeed sufficient to yield a well-defined decision problem and reveal such of the technical details of solving this problem as are necessary for proving the theorem to come.

Letting V_j denote the number of votes received by candidate j from voters other than the focal voter, the focal voter views $(V_1, \ldots, V_K)$ as K-nomial random variables with parameters π and $n - 1$. Thus, each voter's beliefs about the probabilities of various events (such as ties or near ties among the candidates) can be calculated using standard multinomial formulas. The details of this are given in Appendix B. Here, I shall simply assume that the probabilities of all events mentioned are known by the voter.

In particular, the focal voter knows, for all j and k: W_j, the probability that candidate j wins a seat outright (i.e., $|\{h \in \boldsymbol{K}: V_h \geq V_j \text{ and } h \neq j\}| \leq M - 1$), and T_{jk}, the probability that candidates j and k are in a two-way tie for Mth place (i.e., $|\{h \in \boldsymbol{K}: V_h > V_j\}| = M - 1$ and $|\{h \in \boldsymbol{K}: V_h = V_j\}| = 2$). Following Hoffman (1982) and Myerson and Weber (1993), I assume that for large n, the perceived probability of an r-way tie for Mth, $r > 2$, is infinitesimal in comparison to the probability of a two-way tie for Mth. With this assumption, the focal voter's expected utility from abstaining, given u, π, and n, can be approximated with infinitesimal error by

$$\sum_{j=1}^{K} W_j u_j + \sum_{j=1}^{K} \sum_{k>j} T_{jk} \left(\frac{u_j + u_k}{2} \right),$$

where various terms involving r-way ties for Mth, $r > 2$, that are asymptotically negligible in comparison to the largest T_{jk}, have been omitted.

By voting for candidate j, the focal voter can affect his or her utility in two ways (continuing to ignore r-way ties, $r > 2$)—by putting j into a tie with k (yielding a utility increment of $(u_j + u_k)/2 - u_k$) or by breaking a tie between j and k (yielding a utility increment of $u_j - (u_j + u_k)/2$). Assuming (following Hoffman 1982 and Myerson and Weber 1993) that the probability of the event "k is in Mth place, tied with j" equals the probability of the event "k is in Mth place, one vote ahead of j," voting for candidate j rather than abstaining yields an expected utility increment of

$$\xi_j = \sum_{k=1}^{K} T_{jk}(u_j - u_k).$$

Thus,

$$V(u; \pi, n) = \arg\max_{j \in K} \xi_j.$$

APPENDIX B: MULTINOMIAL PROBABILITIES

I shall use the following notation throughout:

R = the set of real numbers;

Q = the set of rational numbers;
Z = the set of nonnegative integers;
ϕ = the standard normal density function;
$\boldsymbol{K} = \{1, \ldots, K\}$, the set of candidates;
$V_1, \ldots, V_K$ are K-nomially distributed random variables with parameters $\pi_1 \geq \pi_2 \geq \ldots \geq \pi_K$ and m;
T_{jk} = probability that candidates j and k are tied for Mth place, $0 < M < K$; and
$\Delta^{K-1} = \{(p_1, \ldots p_K) \in R^K: 0 \leq p_h \leq 1$ for all h & $\Sigma p_h = 1\}$, the simplex in K-space.

I shall explore the asymptotic properties of the tie-probabilities T_{jk}. These probabilities can be expressed as follows. For any set $S \subseteq \boldsymbol{K}$, $|S| = M - 1$, let $\mathrm{up}_j(S)$ denote the event $\cap_{i \in S} V_i > V_j$, $\mathrm{down}_j(S)$ denote the event $\cap_{i \in S} V_i < V_j$, and $E_{jk}(S) = \mathrm{up}_j(S) \cap \mathrm{down}_j(\boldsymbol{K} \backslash S \backslash \{j, k\}) \cap V_j = V_k$. In words, $E_{jk}(S)$ is the event "the $M - 1$ candidates in S finish ahead of j and k, j and k are tied for Mth, and all other candidates finish below j and k." With these definitions,

$$T_{jk} = \sum_{b=1}^{\binom{K-2}{M-1}} \Pr[E_{jk}(S_{jk}^b)],$$

where the sets S_{jk}^b all have $M - 1$ elements and correspond to the different ways of selecting which $M - 1$ candidates—from the $K - 2$ that remain after removing j and k—will have vote totals exceeding j's. The only one of these sets that it is necessary explicitly to define is

$$S_{jk}^1 = \begin{cases} \{1, \ldots, M-1\} & \text{if } j \geq M \\ \{1, \ldots, M\} \backslash \{j\} & \text{if } j < M < k \\ \{1, \ldots, M+1\} \backslash \{j, k\} & \text{if } k \leq M. \end{cases}$$

This is the "natural" ordering of candidates other than j and k. The next lemma is useful in showing that $\Pr[E_{jk}(S_{jk}^1)] \geq \Pr[E_{jk}(S_{jk}^b)]$ for all b.

Lemma 1. *Suppose* $\mathrm{K} > \mathrm{M} + 1$ *and consider any* $\mathrm{A} \subset \mathbf{K} \backslash \{\mathrm{j}, \mathrm{k}\}$ *such that* $|\mathrm{A}| = \mathrm{M} - 1$. *Denote the complement of* A *in* $\mathbf{K} \backslash \{\mathrm{j}, \mathrm{k}\}$ *by* $\mathrm{A^c}$. *Suppose there exists* $\mathrm{d} \in \mathrm{A}$, $\mathrm{u} \in \mathrm{A^c}$, such that $\pi_\mathrm{u} > \pi_\mathrm{d}$, *and define* $\mathrm{B} = \mathrm{A} \backslash \{\mathrm{d}\} \cup \{\mathrm{u}\}$. *Then* $\Pr[\mathrm{E_{jk}(A)}] < \Pr[\mathrm{E_{jk}(B)}]$.

Proof. Note that

$$\Pr[E_{jk}(A)] = \sum_{v \in \bar{E}_{jk}(A)} P(v),$$

where

$$P(v) = m! \prod_{j=1}^{K} \frac{\pi_j^{v_j}}{v_j!}$$

and

$$\bar{E}_{jk}(A) = \left\{ v \in Z^K: \bigcap_{h \in K} V_h = v_h \in E_{jk}(A) \right\}.$$

Consider any term $v \in \bar{E}_{jk}(A)$. The vector v', defined by $v'_h = v_h$ for $h \notin \{d, u\}$, $v'_d = v_u$, and $v'_u = v_d$, is an element of $\bar{E}_{jk}(B)$. Moreover, $P(v)/P(v') = (\pi_d/\pi_u)^{V_d - V_u} \leq 1$ (recall that $\pi_d < \pi_u$ and $v_d > v_u$ since $v \in \bar{E}_{jk}(A)$). QED.

Corollary 1. $Pr[\mathrm{E_{jk}(S_{jk}^1)}] \geq Pr[\mathrm{E_{jk}(S_{jk}^b)}]$ *for all* b.

I omit the proof: it follows directly from Lemma 1. In what follows, I shall abbreviate $E_{jk}(S_{jk}^1)$ by E_{jk}.

Lemma 2. *Let* a, b, *and* c *be elements of* $\mathbf{K}$ *such that either (1)* a < M *and* a < b ≤ M + 1 *and* (b = M + 1 → c ≤ M) *and* $\pi_\mathrm{a} > \pi_\mathrm{b}$ *or (2)* a > M + 1 *and* M ≤ b < a *and* (b = M → c ≥ M) *and* $\pi_\mathrm{a} < \pi_\mathrm{b}$. *Then* $\lim_{\mathrm{m}\to\infty} \mathrm{T_{ca}/T_{cb}} = 0$.

Proof. By Corollary 1, it suffices to show that $\lim_{m\to\infty} \Pr[E_{ca}]/\Pr[E_{cb}] = 0$. The proof proceeds separately for the two cases. In case 1, note that $a < M$ implies $a \in S_{cb}^1$. Note also that $b \leq M$ implies $b \in S_{ca}^1$, since

$$S_{ca}^1 = \begin{cases} \{1, \ldots, M\} \backslash \{a\} & \text{if } M < c \\ \{1, \ldots, M+1\} \backslash \{a, c\} & \text{if } c \leq M. \end{cases}$$

When $b = M + 1$, moreover, $c \leq M$ by hypothesis, and hence $b \in S_{ca}^1$ again. In light of these observations, one can write $E_{ca} = E \cap (V_b > V_c = V_a)$ and $E_{cb} = E \cap (V_a > V_c = V_b)$, where $E = \mathrm{up}_c(S_{ca}^1 \cap S_{cb}^1) \cap \mathrm{down}_c(\boldsymbol{K} \backslash S_{ca}^1 \cap S_{cb}^1 \backslash \{a, b, c\})$.

Let $\pi_{h,a} = E(V_h/m|E_{ca})$ and $\pi_{h,b} = E(V_h/m|E_{cb})$. Note that $\pi_h > 0$ implies $\pi_{h,a} > 0$ and $\pi_{h,b} > 0$. Let $\sigma_{h,a}$ and $\sigma_{h,b}$ denote the conditional standard deviations. Note that they are both finite and that both approach zero as m approaches infinity. Thus the distribution of $(V_1/m, \ldots, V_K/m)$, conditional on E_{ca}, collapses around its mean; that is, for any $\varepsilon > 0$,

$$\lim_{m\to\infty} \Pr[\bigcap_{h=1}^{K} (\pi_{h,a} - \varepsilon < V_h/m < \pi_{h,a} + \varepsilon)|E_{ca}] = 1.$$

A similar statement can be made for the distribution of $(V_1/m, \ldots, V_K/m)$ conditional on E_{cb}.

Let S be the event

$$\bigcap_{h=1}^{K} \frac{V_h}{m} > \min\{\pi_{h,a}, \pi_{h,b}\} - \varepsilon.$$

Note that because

$$S \supseteq \bigcap_{h=1}^{K} (\pi_{h,a} - \varepsilon < V_h/m < \pi_{h,a} + \varepsilon),$$

$$\lim_{m\to\infty} \Pr[S|E_{ca}] = 1.$$

Similarly,

$$\lim_{m\to\infty} \Pr[S|E_{cb}] = 1.$$

In light of these results,

$$\lim_{m\to\infty} \frac{\Pr[E_{ca} \cap S^c]}{\Pr[E_{ca} \cap S]} = \lim_{m\to\infty} \frac{\Pr[E_{ca}](1 - \Pr[S|E_{ca}])}{\Pr[E_{ca}]\Pr[S|E_{ca}]} = 0.$$

Thus since

$$\frac{\Pr[E_{ca}]}{\Pr[E_{cb}]} = \frac{\Pr[E_{ca} \cap S] + \Pr[E_{ca} \cap S^c]}{\Pr[E_{cb} \cap S] + \Pr[E_{cb} \cap S^c]},$$

it suffices to show that

$$\lim_{m \to \infty} \frac{\Pr[E_{ca} \cap S]}{\Pr[E_{cb} \cap S]} = 0.$$

But

$$\frac{\Pr[E_{ca} \cap S]}{\Pr[E_{cb} \cap S]} = \frac{\Pr[E|S] \sum_{p \in \Xi} f_m(p)}{\Pr[E|S] \sum_{p \in \Xi} g_m(p)},$$

where

$$f_m(p) = \Pr[\bigcap_{h \in W} V_h = p_h m | E \cap S]\Pr[V_b > p_c m = V_a | \bigcap_{h \in W} V_h = p_h m],$$

$$g_m(p) = \Pr[\bigcap_{h \in W} V_h = p_h m | E \cap S]\Pr[V_a > p_c m = V_b | \bigcap_{h \in W} V_h = p_h m],$$

$$\Xi = \{p \in \Delta^{K-1} \cap Q^K: \bigcap_{h \in W} V_h = p_h m \subseteq E \cap S \text{ for some } m \in Z\} \text{ and}$$

$W = \boldsymbol{K} \backslash \{a, b\}$.

Fix $p \in \Xi$ and consider the limit as m approaches infinity of $f_m(p)/g_m(p)$. Note that

$$\frac{f_m(p)}{g_m(p)} = \begin{cases} 0/0 & \text{if } p \notin \Xi_m \\ \dfrac{\Pr[V_b > p_c m = V_a | \bigcap_{h \in W} V_h = p_h m]}{\Pr[V_a > p_c m = V_b | \bigcap_{h \in W} V_h = p_h m]} & \text{if } p \in \Xi_m, \end{cases}$$

where $\Xi_m = \{p \in \Xi: p_h m \in Z \text{ for all } h \in W\}$. Note also that

$$\lim_{\substack{m \to \infty \\ m \in \Xi^{-1}(p)}} \frac{f_m(p)}{g_m(p)} = 0,$$

where

$$\Xi^{-1}(p) = \{m \in Z: p \in \Xi_m\}.$$

This last result is derived as follows. Let $v_h = p_h m$ for all h. Conditional on $\cap_{h \in W} V_h = v_h$, V_a and V_b are distributed binomially, with parameters $\pi_a^* = \pi_a/(\pi_a + \pi_b)$, $\pi_b^* = \pi_b/(\pi_b + \pi_a)$, and $m^* = m - \Sigma_{h \in W} v_h$. If $2v_c \geq m^*$, then both probabilities are zero. Conditional on $2v_c < m^*$, we have $(V_a = v_c \to V_b > v_c)$ and $(V_b = v_c \to V_a > v_c)$, since $V_a + V_b = m^* > 2v_c$. Thus (the condition $m \in \Xi^{-1}(p)$ being understood in all that follows)

$$\lim_{m \to \infty} \frac{f_m(p)}{g_m(p)} = \lim_{m \to \infty} \frac{\Pr[v_c = V_a | \bigcap_{h \in W} V_h = v_h]}{\Pr[v_c = V_b | \bigcap_{h \in W} V_h = v_h]}$$

$$= \lim_{m \to \infty} \frac{\phi(v_c - m^* \pi_a^* / \pi_a^*(1 - \pi_a^*))}{\phi(v_c - m^* \pi_b^* / \pi_b^*(1 - \pi_b^*))}$$

$$= \lim_{m \to \infty} \exp \left\{ \frac{(v_c - m^* \pi_b^*)^2 - (v_c - m^* \pi_a^*)^2}{\pi_a^*(1 - \pi_a^*)} \right\}.$$

The last limit equals 0 because

$$\lim_{m \to \infty} v_c^2 - 2v_c m^* \pi_b^* + (m^* \pi_b^*)^2 - (v_c^2 - 2v_c m^* \pi_a^* + (m^* \pi_a^*)^2) = \lim_{m \to \infty} (m^*)^2 (\pi_a^* - \pi_b^*) \left(\frac{2v_c}{m^*} - 1 \right) = -\infty.$$

The last equality follows because $m^* > 2v_c = 2p_c m$ and $p_c > 0$. Thus, $\lim m^* = \infty$, while the second term is a positive constant and the third term a negative constant.

In case 2, $a > M + 1$ implies $a \notin S_{cb}^1$. Moreover, since

$$S_{ca}^1 = \begin{cases} \{1, \ldots, M - 1\} & \text{if } M \leq c \\ \{1, \ldots, M\} \backslash \{c\} & \text{if } c < M, \end{cases}$$

$b \geq M + 1$ and $b = M$ (since this entails $c \geq M$) both imply $b \notin S_{ca}^1$. In light of these observations, one can write $E_{ca} = E \cap (V_c = V_a > V_b)$ and $E_{cb} = E \cap (V_c = V_b > V_a)$, where $E = \text{up}_c(S_{ca}^1 \cap S_{cb}^1) \cap \text{down}_c(\boldsymbol{K} \backslash S_{ca}^1 \cap S_{cb}^1 \backslash \{a, b, c\})$.

Given these definitions, the proof here proceeds identically to that in the first case if one substitutes the string '$p_c m = V_a > V_b$' for every occurrence of the string '$V_b > p_c m = V_a$' and the string '$p_c m = V_b > V_a$' for every occurrence of the string '$V_a > p_c m = V_b$'. Somewhat larger changes are needed when one examines the limit of the ratio $f_m(p)/g_m(p)$, where, in this case,

$$\frac{f_m(p)}{g_m(p)} = \begin{cases} 0/0 & \text{if } p \notin \Xi_m \\ \dfrac{\Pr[p_c m = V_a > V_b | \lim_{h \in W} V_h = p_h m]}{\Pr[p_c m = V_b > V_a | \lim_{h \in W} V_h = p_h m]} & \text{if } p \in \Xi_m. \end{cases}$$

Conditional on $\cap_{h \in W} V_h = v_h$, V_a and V_b are distributed binomially, with parameters $\pi_a^* = \pi_a/(\pi_a + \pi_b)$, $\pi_b^* = \pi_b/(\pi_b + \pi_a)$, and $m^* = m - \Sigma_{h \in W} V_h$. If $m^* < v_c$ or $m^* \geq 2v_c$, then both probabilities are zero. Conditional on $v_c \leq m^* < 2v_c$, we have $(V_a = v_c \rightarrow V_a > V_b)$ and $(V_b = v_c \rightarrow V_b > V_a)$. Thus (the condition $m \in \Xi^{-1}(p)$ being understood in all that follows),

$$\lim_{m \to \infty} \frac{f_m(p)}{g_m(p)} = \lim_{m \to \infty} \frac{\Pr[v_c = V_a | \lim_{h \in W} V_h = v_h]}{\Pr[v_c = V_b | \lim_{h \in W} V_h = v_h]}$$

$$= \lim_{m \to \infty} \frac{\phi(v_c - m^*\pi_a^*/\pi_a^*(1 - \pi_a^*))}{\phi(v_c - m^*\pi_b^*/\pi_b^*(1 - \pi_b^*))}$$

$$= \lim_{m \to \infty} \exp\left\{\frac{(v_c - m^*\pi_b^*)^2 - (v_c - m^*\pi_a^*)^2}{\pi_a^*(1 - \pi_a^*)}\right\}.$$

The last limit equals zero because

$$\lim_{m \to \infty} v_c^2 - 2v_c m^* \pi_b^* + (m^*\pi_b^*)^2 - (v_c^2 - 2v_c m^* \pi_a^* + (m^*\pi_a^*)^2) = \lim_{m \to \infty} (m^*)^2(\pi_a^* - \pi_b^*)(2v_c m^* - 1) = -\infty.$$

The last equality follows because $m^* < 2v_c$ (so the last term is a positive constant), the second term is a negative constant, and $m^* \geq p_c m > 0$ (so $\lim m^* = \infty$). QED.

LEMMA 3. *(a) If* $\pi_1 > \pi_M$, *then* $\lim_{m\to\infty} T_{1j} / T_{M,M+1} = 0$ *for all* j. *(b) If* $0 < \pi_j < \pi_{M+1}$ *for some* $j > M + 1$, *then* $\lim_{m\to\infty} T_{jk} / T_{M,M+1} = 0$ *for all* k.

Proof Part 1. By Corollary 1, it suffices to show that $\lim_{m\to\infty} \Pr[E_{1j}]/\Pr[E_{M,M+1}] = 0$ for all j. And this follows because (1) $\lim_{m\to\infty} \Pr[E_{1j}]/\Pr[E_{1M}] = 0$ for all $j < M$ (by Lemma 2, case 1); (2) $\lim_{m\to\infty} \Pr[E_{1j}]/\Pr[E_{1,M+1}] = 0$ for all $j > M + 1$ (by Lemma 2, case 2); (3) $\lim_{m\to\infty} \Pr[E_{1M}]/\Pr[E_{M,M+1}] = 0$ (by Lemma 2, case 1); and (4) $\lim_{m\to\infty} \Pr[E_{1,M+1}]/\Pr[E_{M,M+1}] = 0$ (by Lemma 2, case 1).

Proof Part 2. By Corollary 1, it suffices to show that $\lim_{m\to\infty} \Pr[E_{jk}]/\Pr[E_{M,M+1}] = 0$ for all k. And this follows because (1) $\lim_{m\to\infty} \Pr[E_{jk}]/\Pr[E_{j,M+1}] = 0$ for all $k > M + 1$ (by Lemma 2, case 2); (2) $\lim_{m\to\infty} \Pr[E_{jk}]/\Pr[E_{jM}] = 0$ for all $k < M$ (by Lemma 2, case 1); (3) $\lim_{m\to\infty} \Pr[E_{j,M+1}]/\Pr[E_{M,M+1}] = 0$ (by Lemma 2, case 2); and (4) $\lim_{m\to\infty} \Pr[E_{jM}]/\Pr[E_{M,M+1}] = 0$ (by Lemma 2, case 2). QED.

Proof of Theorem 1. Suppose $\pi_1 > \pi_M$ and consider a voter of type u. Let $X_j = \{k \in K: \pi_k = \pi_j\}$ and call any candidate in $X_M \cup X_{M+1}$ "marginal." Let h be the marginal candidate whom the voter ranks highest: $u_h > u_j$ for all $j \in X_M \cup X_{M+1} \backslash \{h\}$. (Such a candidate exists because all voters have strict preferences among candidates.)

Now consider ξ_h^n, the expected utility increment of voting for h rather than abstaining, when the electorate is of size n. From Appendix A, $\xi_h^n = \Sigma_{k=1}^K T_{hk}^n(u_h - u_k)$, where T_{hk}^n *is the probability of a two-way tie for* Mth between h and k, given an electorate of size n (and given π). Note that $\xi_h^n > 0$ for all n, because (1) in the limit, ties between marginal candidates dominate those involving nonmarginal candidates, by Lemma 2 and (2) $u_h - u_k > 0$ for all $k \in X_M \cup X_{M+1} \backslash \{h\}$. Note also that $\lim \xi_1^n / \xi_h^n = 0$, by Lemma 3. (One can assume without loss of generality that $h \leq M + 1$, since $h \in X_M \cup X_{M+1}$.) Thus, in the limit, no voter will vote for candidate 1, given expectations π, which contradicts the expectation that $\pi_1 > 0$. This shows that for large enough n, π is not a rational expectation. But there may be rational expectations arbitrarily close to π.

In order to show that one cannot find any rational expectations "close" to π, consider a sequence $\{\pi^n\}$ of rational expectations that converges to π. The discussion leading up to Lemma 1, and the proof of that lemma, show that

$$T_{hk}^n = \sum_{b=1}^{\binom{M-1}{K-2}} \sum_{v \in \bar{E}_{hk}(S_{hk}^b)} (n-1)! \prod_{j=1}^{K} \frac{\pi_j^{v_j}}{v_j!},$$

where

$$\bar{E}_{hk}(S_{hk}^b) = \{v \in Z^K: \lim_{h \in K} V_h = v_h \in E_{hk}(S_{hk}^b)\}.$$

Thus T_{hk}^n is a continuous function of π, and so is ξ_h^n. Let $\xi_h^n(\pi^n)$ be the expression resulting when one substitutes π^n for π in the expression for ξ_h^n. The continuity of ξ_h^n in π, together with $\pi^n \rightarrow \pi$ and $\xi_h^n > 0$ for all n, implies $\xi_h^n(\pi^n) > 0$ for all n sufficiently large. The ratio ξ_1^n/ξ_h^n is continuous in π, over the set of all π such that ξ_h^n is not zero. Thus $\lim \xi_1^n(\pi^n)/\xi_h^n(\pi^n) = \lim \xi_1^n/\xi_h^n = 0$, the latter equality having been established above; that is, in sufficiently large electorates with expectations sufficiently close to π, no voter will vote for candidate 1, contrary to the assumption that $\pi_1 > 0$. A similar proof works for the other part of the theorem. QED.

Notes

I thank Skip Lupia, Barry Nalebuff, Tom Palfrey, and especially Roger Myerson for their helpful comments. Thanks also to Jonathan Katz for suggesting and coding the Silverman bimodality tests conducted in the fourth section. This work was supported by the National Science Foundation under grant number SES-9208753.

1. The Japanese and Taiwanese require that a candidate get above a legally defined minimum vote share in order to win a seat. The seat allocation rule, in other words, is not pure plurality. Nonetheless, this provision is of negligible practical importance.

2. The adjective *strategic* in *strategic voting equilibria* is really superfluous in that agents in the model conform to the usual postulate of rationality and are not distinctly more strategic than agents in any other rational choice model. The term *strategic voting* is well established in the literature, however, and I shall follow conventional usage.

3. Myerson and Weber (1993) suggest that considerations of dynamic consistency or persistence may eliminate the non-Duvergerian equilibria but do not demonstrate this.

4. Myerson and Weber (1993) show that neither one of these assumptions is crucial.

5. Formally, the assumption is this. Let each voter order the candidates from largest π_{ij} to smallest. Let $M(i)$ be the candidate ranked Mth in i's ordering and $M + 1(i)$ be the candidate ranked $(M + 1)$st. Let $\text{Mid}(i) = \{j: \pi_{ij} = \pi_{iM(i)} \text{ or } \pi_{ij} = \pi_{iM+1(i)}\}$, $\text{Lead}(i) = \{j: \pi_{ij} > \pi_{iM(i)}\}$, and $\text{Trail}(i) = K\backslash\text{Mid}(i)\backslash\text{Lead}(i)$. Then, for all voters i and h, $\text{Lead}(i) = \text{Lead}(h)$ and $\text{Trail}(i) = \text{Trail}(h)$.

6. Because there are no atoms in the type distribution F, nonsingleton $V(u; \pi, n)$s do not require special handling.

7. I thank Roger Myerson and Barry Nalebuff for greatly clarifying what this theorem ought to say.

8. Trailing candidates find allies in their attempts to avoid the logic of the wasted vote in front-runners who expect a net loss should the trailing candidacy go down the tubes (e.g., recall Ronald Reagan's support of John Anderson's candidacy in 1980) and foes in front-runners who expect a net gain of support (recall Jimmy Carter's persistent reminders to voters not to waste their vote on Anderson).

9. The data used in this analysis are from Steven Reed's compendium *Japan Election Data: The House of Representatives 1947–1990* (Ann Arbor: Center for Japanese Studies, 1992).

10. Note that there is an artifactual expectation of precisely the opposite character. Looking at the mathematical definitions of EXCESS and MARGIN given in the notes to Table 1, one sees that V_M, the vote percentage of the Mth-place candidate, appears with a negative sign in the expression for EXCESS and with a positive sign in the expression for MARGIN. Thus one expects an artifactual negative relationship. If the theoretically expected positive relationship nonetheless appears, one can be more confident that it corresponds to some real political effect.

References

Austen-Smith, David. 1987. "Sophisticated Sincerity: Voting over Endogenous Agendas." *American Political Science Review* 81:1323–30.

Banks, Jeffrey. 1985. "Sophisticated Voting Outcomes and Agenda Control." *Social Choice and Welfare* 1:295–306.

Cox, Gary W. 1984. "Strategic Electoral Choice in Multi-Member Districts: Approval Voting in Practice?" *American Journal of Political Science* 28:722–38.

Cox, Gary W. 1987. "Duverger's Law and Strategic Voting." University of California, San Diego. Typescript.

Cox, Gary W. 1991. "Duverger's Law: An Interpretive Review of Research." University of California, San Diego. Typescript.

Cox, Gary W., and Michael Munger. 1989. "Closeness, Expenditures, and Turnout in the 1982 U.S. House Elections." *American Political Science Review* 83:217–231.

Cox, Gary W., and Emerson Niou. 1994. "Seat Bonuses under the Single Non-Transferable Vote System: Evidence from Japan and Taiwan." *Comparative Politics* 26:221–236.

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

Duverger, Maurice. 1955. *Political Parties*. New York: Wiley.

Farquharson, Robin. 1969. *Theory of Voting*. New Haven: Yale University Press.

Green, Donald P., and Ian Shapiro. 1993. "Pathologies of Rational Choice Theory: A Critique of Applications in Political Science." Institution for Social and Policy Studies Working Paper No. 1043. Yale University.

Hoffman, Dale T. 1982. "A Model for Strategic Voting." *SIAM Journal of Applied Mathematics* 42:751–61.

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

Ledyard, John. 1984. "The Pure Theoryy of Two Candidate Elections." *Public Choice* 44:7–41.

McKelvey, Richard, and Peter Ordeshook. 1972. "A General Theory of the Calculus of Voting." In *Mathematical Applications in Political Science*, vol. 6, ed. J. F. Herndon and J. L. Bernd. Charlottesville: University Press of Virginia.

McKelvey, Richard, and Richard Niemi. 1978. "A Multistage Representation of Sophisticated Voting for Binary Procedures." *Journal of Economic Theory* 18:1–22.

Meehl, P. E. 1977. "The Selfish Voter Paradox and the Thrown-away Vote Argument." *American Political Science Review* 61:11–30.

Miller, Nicholas. 1980. "A New Solution Set for Tournaments and Majority Voting." *American Journal of Political Science* 24:68–96.

Myerson, Roger, and Robert Weber. 1993. "A Theory of Voting Equilibria." *American Political Science Review* 87:102–114.

Niemi, Richard. 1984. "The Problem of Strategic Voting under Approval Voting." *American Political Science Review* 78:952–58.

Ordeshook, Peter, and Thomas Schwartz. 1987. "Agendas and the Control of Political Outcomes." *American Political Science Review* 81:179–99.

Palfrey, Thomas. 1989. "A Mathematical Proof of Duverger's Law." In *Models of Strategic Choice in Politics*, ed. Peter C. Ordeshook. Ann Arbor: University of Michigan Press.

Reed, Steven R. 1991. "Structure and Behaviour: Extending Duverger's Law to the Japanese Case." *British Journal of Political Science* 29:335–56.

Riker, William H. 1982. "The Two-Party System and Duverger's Law: An Essay on the History of Political Science." *American Political Science Review* 76:753–66.

Shepsle, Kenneth, and Barry Weingast. 1984. "Uncovered Sets and Sophisticated Voting Outcomes with Implications for Agenda Control." *American Journal of Political Science* 28:49–74.

Silverman, B. W. 1981. "Using Kernel Density Estimates to Investigate Multimodality." *Journal of the Royal Statistical Society*, ser. B 43:97–99.

Gary W. Cox is Professor of Political Science, University of California—San Diego, La Jolla, CA 92093-0521.