
*Strategic voting in single-member
single-ballot systems*

“The evidence renders it undeniable that a large amount of sophisticated voting occurs – mostly to the disadvantage of the third parties nationally – so that the force of Duverger’s psychological factor must be considerable.”

William H. Riker (1982:764).

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. In this chapter, I 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 thirty years and focuses on the behavior of legislators (e.g., Farquharson 1969; McKelvey and Niemi 1978; Miller 1980; Shepsle and Weingast 1984; Banks 1985; Austen-Smith 1987; Ordeshook and Schwartz 1987). This chapter focuses on the other, less well-trodden, branch of research into strategic voting, that dealing with the behavior of voters in mass elections.

There are of course many ways to conduct a mass election. This chapter deals in particular with elections in electoral districts that satisfy the following criteria: (1) There is one seat to be filled in the district; and (2) there is only one round of voting, after which the victor is decided. There are many electoral systems that satisfy these criteria: the Anglo-American first-past-the-post system; the Australian alternative vote system; the approval voting system; and so on. Not all of these systems are currently used in national elections, of course. For the most part, I shall focus on those that are.

The purpose of this chapter is primarily to specify the theoretical and institutional conditions under which Duverger’s Law holds at the local

Strategic voting

level. In one sense, my findings are largely negative: If one changes any of several institutional features that define ordinary plurality rule (e.g., by allowing fusion candidacies, holding all else constant), then the strategic voting incentives that push toward local bipartism dissipate substantially; if one investigates the theoretical conditions that are necessary to generate a strong local bipartism result, even *given* the right institutional context (i.e., ordinary plurality rule), they appear fairly stringent; and if one investigates the empirical evidence at the district level in countries that use ordinary plurality rule, one finds plenty of cases where more than two candidates enter and receive substantial vote shares, contrary to Duvergerian expectations. On the other hand, my findings are by no means all negative: The institutional conditions can be met (just use ordinary plurality rule), the theoretical conditions are plausible in certain situations, and when the theoretical conditions are approximated in the real world, one finds ample evidence consistent with strategic behavior, as suggested by the quote from William H. Riker that heads this chapter.

The layout of the chapter is as follows. The first section reviews the previous formal literature on strategic voting (leaving the vast informal literature largely untouched). The next four sections consider strategic voting in three particular electoral systems – ordinary plurality rule (Sections 4.2 and 4.3), plurality rule with fusion candidacies allowed (4.4), and the alternative vote (4.5). Section 4.6 concludes.

I should stress before proceeding further that this chapter, as well as the succeeding chapters in this part of the book, all take a post-entry district-level perspective. Duverger's Law claims that use of single-member districts operating under plurality rule will lead to bipartism at the *national* level. To build up to this national-level conclusion, however, Duverger starts at the district level, arguing that strategic voting (and strategic entry) should produce *local* bipartism in each district. How a series of potentially disconnected local two-party systems might cumulate to a national two-party system is a topic for a later chapter (10). In this chapter, I consider the logic of the foundational claim that use of ordinary plurality in single-member districts will lead to local bipartism.

In assessing this claim, it is conventional to note that ordinary plurality mechanically underrepresents small parties at the district level (because the winner takes all) and thereby stimulates two species of strategic adaptation: strategic voting by citizens eager not to waste their votes; and strategic withdrawals by politicians eager not to waste their effort and resources (cf. Blais and Carty 1991:83). In this part of the book, I focus exclusively on the theory and practice of strategic voting (leaving the theory and practice of strategic entry/withdrawal to the next part).

Strategic voting

4.1 PREVIOUS RESEARCH

Formal mathematical study of strategic voting in the last twenty 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, for simple plurality elections, roughly as follows: Some voter, whose favorite candidate has a poor chance of winning, notices that she has a preference between the top two candidates; she then rationally decides to vote for the most preferred of these top two competitors rather than for her overall favorite, because the latter vote has a much smaller chance of actually affecting the outcome than the former. What the decision-theoretic approach adds to common sense is not just greater precision about the assumptions implicit in such reasoning (for example, it is not the probability of victory that matters directly, but the probability of certain ties and near-ties) but also greater generality: The basic model has been extended to illuminate strategic behavior in Borda elections (Ludwin 1978; Dummett 1984), in multimember districts (Cox 1984), in approval voting elections (Niemi 1984), and in a variety of other electoral systems (Hoffman 1982; Dummett 1984; Gutowski and Georges 1993).

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 front-runners). 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 front-runners. 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 first addressed these questions in the context of a model in which three candidates compete for a single seat under simple plurality (Cox 1987b). The key assumptions of the analysis were that all voters are short-term instrumentally rational (i.e., they care about whom they vote for only insofar as this affects the outcome of the current election), that voters have incomplete information about each other's preferences over

Strategic voting

outcomes, and that all voters have “rational” expectations (on which, more later). 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). The logic of this result is worth sketching, as I shall refer to it again.

Strategic voting in a simple plurality election means voting for a lower-ranked candidate that one believes is stronger, rather than for a higher-ranked candidate that one believes is weaker. The rational expectations condition implies that voter beliefs about which candidates are stronger and weaker will be generally correct. Thus, strategic voting will generally transfer votes from objectively weaker (vote-poorer) to objectively stronger (vote-richer) candidates, with the necessary result that the “effective number of parties” – a measure which is smaller the more concentrated the distribution of votes is – will decline. There are many other electoral systems, it should be noted, for which this result does not hold (see Chapter 7).

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). Duvergerian equilibria are so named because they gibe with Duverger’s expectations that simple plurality will promote bipartisanship. The intuition behind the non-Duvergerian equilibria is roughly as follows. Suppose two leftist candidates (say, Charles Goodell and Richard Ottinger) and one rightist (say, James Buckley) are competing for a single post (one of the U.S. Senate seats for New York). 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.¹

Myerson and Weber (1993) advance a model of voting equilibria applicable in a wide range of single-winner electoral systems – not just simple plurality rule but also approval voting, Borda’s method of points, and many other systems as well. Their approach is more general in that, where Cox and Palfrey assume a particular (multinomial) model of voter probability beliefs, Myerson and Weber merely require that these beliefs satisfy a fairly general requirement (the “ordering condition,” whereby 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

¹On the theoretical status of the non-Duvergerian equilibria, see Myerson and Weber (1993:106) and Fey (1995).

Strategic voting

place first or second in the poll). On the other hand, the Cox-Palfrey approach has the virtue of deriving the ordering condition endogenously as a consequence of more primitive assumptions.

4.2 STRATEGIC VOTING UNDER SIMPLE PLURALITY WITHOUT FUSION

A theoretical model

In this section I sketch out a model of strategic voting in simple plurality elections, based on Cox (1994).² Although a few mathematical symbols creep into the text in this and the next section, technical details are left to the footnotes, and the discussion returns to an “English-only” status thereafter.

Imagine K candidates competing for a single seat, with the candidate placing highest in the poll winning. Each voter casts a single exclusive vote and can be characterized by her *preferences* among the candidates, her *beliefs* about the candidate preferences of other voters, and her *expectations* about the likely outcome of the election.

Preferences. I assume that each voter i cares about which candidate wins the election, these preferences being formally represented by a von Neumann-Morgenstern utility function u_i .³ Following standard usage in game theory, I shall sometimes refer to u_i as voter i 's *type*.

Beliefs. No voter knows the candidate preferences of other voters with certainty, but each does have beliefs about how frequently the various different types of voter crop up in the electorate as a whole. Formally, these beliefs are encapsulated in a cumulative distribution function F_i .⁴

Expectations. Voters also have beliefs, or *expectations*, about how well each candidate is likely to do in the upcoming election. These expectations are formalized as a vector $\pi_i = (\pi_{i1}, \dots, \pi_{iK})$, where π_{ij} denotes the proportion of the electorate that i expects will vote for j . Given preferences (u_i) , expectations (π_i) , and knowledge of the number of voters (n) , voter i faces a standard decision problem, the details of which are

²This model is essentially the same as that of Palfrey (1989), although the method of proof differs. It is also closely related to the work of Myerson and Weber (1993).

³Voters' utilities can be rescaled in the standard fashion so that victory for the voter's most-preferred candidate yields a utility of 1, while victory for her least-preferred candidate yields a utility of 0. After this rescaling, voter i 's preferences (or voter i 's *type*) can be described by the vector $u_i = (u_{i1}, \dots, u_{iK})$, an element in the set $U = \{(u_1, \dots, u_K) : \max\{u_j\}=1 \ \& \ \min\{u_j\}=0 \ \& \ u_j=u_k \ \text{only if } j=k\}$.

⁴Given F_i , which is defined over the set U , one can define a distribution over U^n , assuming independence. An alternative approach is to make assumptions directly about the distribution over U^n (over profiles) instead of over U (over individuals).

Strategic voting

run through in Cox (1994). 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 $V(u_i; \pi_i, n)$. $V(u_i; \pi_i, n)$ is simply "the optimal vote for a voter with preferences u_i and expectations π_i " (although in some instances the voter may be indifferent between two or more vote choices, in which case we would need to talk of "the set of optimal votes").

The model is completed with two further assumptions whose joint effect is to restrict the nature and consistency of voter beliefs (about other voters' preferences) and expectations (about how well each candidate will do). First, I assume that $F_i = F$ for all i . In other words, all voters share a common view of the distribution of voter preferences in the electorate.⁵ Second, I assume that voters' expectations are publicly generated – by, for example, 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 . Both of these assumptions can be replaced with weaker ones, under which voters do not agree exactly on how preferences are distributed in the electorate or on what share of the vote each candidate will likely get, without destroying the key result to come.

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 voters' knowledge of the distribution F of voter preferences. Suppose, to take a three-candidate example, that some voter thought π equaled $(1/3, 1/3, 1/3)$, so that a randomly selected voter was equally likely to vote for any of the candidates. This expectation is not consistent with a distribution of voter preferences in which the proportion of voters ranking candidate 1 last exceeds $2/3$. The reason is that voting for 1 is a dominated strategy for voters who rank him last; thus, even if every voter not ranking 1 last intends to vote for him, this still falls short of one-third of the electorate, hence short of the proportion expected under π .

Considerations such as these motivate imposing the following "rational expectations" condition on voter beliefs:

Rational expectations condition: The expectations π are rational with respect to the beliefs F if an electorate whose preferences were in fact distributed according to F , all voting optimally in light of π

⁵Another technical assumption employed about F is that its support set is U . That is, each voter entertains the possibility that there are some voters of any given preference type in the electorate, although they may assign a very low probability to some (or even most) such types.

Strategic voting in single-member single-ballot systems
(i.e., casting a ballot in $V(u; \pi, n)$), would in fact produce expected vote shares for the candidates identical to π .

The equilibrium conditions for the model are then two. First, every voter votes so as to maximize her expected utility, given expectations π (and n); that is, every voter of type u votes for candidate $V(u; \pi, n)$. Second, the expectations π satisfy the rational expectations condition.

Voting equilibria and wasted votes

What are the equilibria of the model just sketched? Relabel the candidates, if necessary, so that their labels correspond to their expected rank of finish, i.e., so that $\pi_1 \geq \pi_2 \geq \dots \geq \pi_K$. Note that with this relabeling one can reasonably assume $\pi_2 > 0$: The candidate expected to place second has a positive expected vote share. Given a distribution F of voter types, I shall say that the expectations π are a limit of rational expectations if and only if arbitrarily large electorates can have rational expectations that are arbitrarily close to π . The point of considering “large electorates” is that expectations in the model become arbitrarily precise in the limit, so that there is a simple relationship between the expected order of finish of the candidates and their probabilities of winning seats. The main result is presented in the following theorem and its corollary.

Theorem 1: Suppose that $0 < \pi_j < \pi_2$ for some $j > 2$. Then π is not a limit of rational expectations.

Proof: See Cox 1994.

The basic logic of the proof is this: If $0 < \pi_j < \pi_2$, 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 first yields a higher expected utility than voting for j . A direct consequence of Theorem 1 is:

Corollary 1: If π is a limit of rational expectations, then $\pi_j \in \{0, \pi_2\}$ for all $j > 2$.

The corollary divides equilibria into two classes: (1) Duvergerian equilibria, with two vote-getting candidates; and (2) non-Duvergerian equilibria, with more than two vote-getting candidates. The Duvergerian equilibria entail a single runner-up, all other candidates being reduced to near-zero support. The non-Duvergerian equilibria entail two or more runners-up, whose nearly identical expected vote totals prevent any being winnowed out from the field of viable candidates.

The intuitive motivation for the results just presented is as follows. Imagine a particular expected order of finish between $K = 3$ candidates:

Strategic voting

Candidate 1 is expected to finish first, 2 second, and 3 third. Nothing is said about how far ahead of candidate $j + 1$ candidate j is. It might be a small proportional difference, or a large one. If it is small, and the electorate is small, then it is not hard to imagine that $j + 1$ might in fact finish ahead of j . In particular, if 3 is close to 2, then the chances of 1 and 2 tying for first may not be much greater than the chances of 1 and 3 tying for first. As the electorate grows, however, and assuming that the distribution of vote shares collapses around its mean (i.e., π), it becomes less and less plausible that 3 might overhaul 2 and compete with 1 for the seat.⁶ Thus, votes for 3 become less and less attractive from the point of view of affecting the outcome, relative to votes for 1 and 2, with the consequence that all short-term instrumental voters desert 3 for either 1 or 2.

Key assumptions

The model discussed in the previous sections embodies one set of assumptions that are sufficient to produce pure local bipartism. Ignoring the non-Duvergerian equilibria for the moment, the Duvergerian equilibria yield a strong version of Duverger's Law: All third parties are reduced to zero support, utterly devastated by strategic voting. This is local bipartism with a vengeance.

What are the assumptions in the model necessary to produce this result? I shall mention the four that seem most important.

Note first that if preferences are not strict, then the reduction of trailing candidates (those expected to place third or lower) to zero is not necessary. Candidate 3's supporters will never desert him if they rank 1 and 2 *equally*. For then there is nothing to choose between the front-runners; any voter who most prefers 3 has a dominant strategy actually to vote for 3. Allowing for the possibility of indifference, one would have to modify the conclusion of Theorem 1: Trailing candidates would be reduced, not to zero support, but to their "hard-core" support (consisting of all those who viewed the front-runners as equally bad alternatives). This is not a terribly important caveat if there are not many voters who are (nearly) indifferent between two or more candidates. But there may be situations in which fairly large numbers of voters do feel intensely about their first choices and relatively weakly about the difference between their second and third choices, in which case the effect of expectations may be relatively small relative to that of preferences.

⁶Even if the probability q_{12} that 1 and 2 tie tends to zero, the probability q_{13} that 1 and 3 tie is so small that q_{13}/q_{12} tends to zero. Put another way, even given the unlikely event that a tie for first occurs, the probability that this tie is between 1 and 3 tends to zero as the electorate increases.

Strategic voting in single-member single-ballot systems

This is the pattern that Blais and Nadeau (N.d.) find in some Canadian elections.

Another technical assumption that is necessary to reduce trailing candidates to zero is that all types of voters are represented in the electorate (the support set of F is U). If one dispenses with this assumption, then it may be that a particular candidate has such an advantage in terms of the distribution of voter preferences that he will win with certainty. Suppose, for example, that there is a unidimensional policy space along which the parties are arrayed; a large centrist party preferred by, say, 45% of the electorate; and a smattering of small parties to the left and right. In this case, assuming that a party's spatial position captures everything about it that voters value and that all voters have single-peaked preferences, it will be common knowledge that the centrist party can defeat any other single party in a pairwise competition. It will not be politically feasible to construct an ends-against-the-center coalition, and the victory of the centrist will be certain. Accordingly, there will be no reason for a voter to desert his or her favorite leftist or rightist party: Small parties will continue, and strategic voting will be minimal, in the face of a Condorcet winner. This is essentially Riker's analysis of the Indian case, with the Congress Party playing the role of the centrist (Riker 1976).

A more pragmatic take on the same point goes as follows. The more obvious it is that a particular candidate is going to win, the less pressure there is to vote strategically. The less obvious it is who will win, the more pressure to vote effectively rather than expressively.

A third, more substantive, assumption that is necessary to preserve the strong local bipartism result derived in Theorem 1 is that all voters are short-term instrumentally rational. This assumption excludes voters who take a long-term, albeit still instrumental, viewpoint: voters, for example, who seek to affect the outcome of future elections by demonstrating stubbornness in this election. (Supporters of the Prohibition Party in the turn-of-the-century United States may have believed that by demonstrating a willingness to incur the cost of a bad outcome this time, they could convince their most likely major-party partner to adopt their viewpoint on liquor. Similarly, perhaps those who voted for the richest of the three candidates in the 1992 U.S. presidential election can be characterized as "waiting for Perot," rather than as miscalculating the then-relevant electoral probabilities. Such speculations cast an interesting light on the notion of electoral realignment but are not pursued here.) Assuming short-term instrumental rationality also excludes voters who derive a direct consumption value from the act of voting for one or another candidate: voters, for example, who use their vote to affirm allegiance to a political cause. Introducing voters who are not short-term instrumentally rational into the model modifies the result of Theorem 1 roughly as fol-

Strategic voting

lows: Trailing candidates are deserted, not by all voters, but by all short-term instrumentally rational voters (cf. Cox 1994). The more short-term instrumentally rational voters there are, then, the more closely does the theoretically predicted result approximate that of the baseline model.

A fourth condition necessary to generate pure local bipartism is that the identity of trailing and front-running candidates is common knowledge. The extent to which this knowledge is public keeps all instrumental voters on the same page of the playbook: They *all* desert the (publicly identified) trailing candidates in order to focus on the (publicly identified) front-running candidates. (There are several assumptions in the model that contribute to this certainty and consensus on the part of voters regarding candidate chances but two are particularly important: first, that voters' expectations are rational; second, that in the limit voters can be virtually certain about the candidates' order of finish.)

One might argue for the reasonableness of the common knowledge assumption by noting the self-fulfilling character of voter expectations. If every voter *believes* that candidate *j* is out of the running, then he will in fact *be* out of the running. Moreover, if some voters, who previously intended to vote for *j*, come to believe that he is behind, they will desert him, thereby making it more likely that he 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 game will coordinate on one of the two pure-strategy Nash equilibria.

If who trails 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 (cf. Johnston et al. 1992:197–211). It has to be free information because rationally ignorant voters will not exert any effort in determining 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

Strategic voting in single-member single-ballot systems

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) which reveals some candidates to be clearly trailing the others, and 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.

One reasonable reaction to the list of conditions necessary to produce pure local bipartism might be that they illuminate the *limits* to Duverger's reasoning. That there are such limits Duverger himself emphasized: that is why he stated his law as only a tendency. The advantage of the formal model is that it specifies some of the limitations. In particular, the model suggests that failures to achieve the drastic reduction in third party vote totals predicted by Theorem 1 can flow from (1) the presence of voters who are not short-term instrumentally rational; (2) lack of public information about voter preferences and vote intentions (hence about which candidates are likely to be “out of the running”); (3) public belief that a particular candidate will win with certainty; or (4) the presence of many voters who care intensely about their first choice and are nearly indifferent between their second and lower choices.

Although quite a few assumptions are needed to generate a pure local bipartism result, it should be noted that much less is needed to generate appropriate comparative statics results. To generate a tendency toward bipartism it is sufficient, for example, to posit (1) short-term instrumentally rational voters; (2) reasonably accurate and publicly available information on candidate standings (π); and (3) myopic (“price-taking”) adjustment. Such a dynamic adjustment model will converge to a Duvergerian equilibrium. Consider, for example, a situation in which the percentage of voters ranking Candidate 1 first is 36%, the percentage ranking 2 first is also 36%, and the percentage ranking 3 first is 28%. A sequence of r random-sample polls is taken, each with a margin of error of $\pm 1\%$. If all voters answer the first poll sincerely and then respond truthfully regarding their current vote intentions, it will rapidly become evident that the chance of 3 tying for first is small relative to that of 1 and 2 tying for first. Thus, 3's least-committed supporters – for whom 1 or 2 are good substitutes –

Strategic voting

will desert him. The next poll will reflect this desertion and lower the chance of 3 tying for first even more, and so on (cf. Fey 1995).⁷

4.3 STRATEGIC VOTING AS AN EXPLANATION OF REAL-WORLD PHENOMENA

In this section, I consider the empirical usefulness of the results just sketched. There is no question that short-term 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, hence to doubt the result that strategic voting will devastate third parties.

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. I have suggested what some of the noise to be added might be above. Even if adding noise, say in the form of noninstrumental voters or of voters whose expectations are inconsistent, can in principle produce predictions that tend toward local bipartism without going all the way, there is still interest in two questions: First, do real-world data conform sufficiently closely to the model's predicted equilibria so that one might believe that a model essentially similar to this one (just adding noise) might tally with real-world patterns? Second, even if the real world conforms to stylized versions of the model's equilibria, are there other explanations that predict the same patterns? I shall examine each of these topics – empirical patterns and alternative explanations – in turn.

Empirical patterns: The literature

The main pattern that the model predicts is the strategic desertion of trailing candidates by their instrumental supporters. Empirically, there is substantial evidence in the literature that real voters do vote strategically in simple plurality contests for legislative office, whether one talks of elections to the German *Bundestag* (see below), the British House of Commons (see below), the Liverpool City Council (Laver 1987), the Canadian House of Commons (Blais, Renaut, and Desrosiers 1974; Black 1978, 1980; Bowler and Lanoue 1992; Blais and Nadeau N.d.), or

⁷Even if one allows voters to answer polls strategically, this should not change the outcome much. What might change the result is if the margin of error in the poll were large relative to the difference in support between candidates.

Strategic voting in single-member single-ballot systems

the New Zealand House of Commons (Vowles and Aimer 1993:25, 157; Catt 1991; Rydon 1989:137; Levine and Roberts 1991).⁸

To give an idea of what is, and is not, in the literature, I shall consider the German and British evidence in more detail. One might question whether the German evidence really belongs in the simple plurality column. It is true that voters in each constituency possess a single exclusive vote and that plurality rule determines the winner. But Germans also have a second vote which they may cast for a list within their *Land*, and it is the list votes that determine how many seats each party will receive. So why would German voters care who won in their district? One way to think of it is in terms of the *Überhangmandat* clause, whereby parties that win more constituency seats than their list votes would be entitled to nonetheless keep their “extra” seats. In light of this rule, electing another Christian Democrat as a constituency candidate may make sense to a Free Democrat (FDP) voter who detests the major alternative, the Social Democrats (SPD). Readers who believe that this explanation demands too much of the German voter may find another idea more plausible: that the identity of the local representative is valued in itself, above and beyond the balance of party forces in the *Bundestag*. Either way, strategic voting in German constituencies should be similar to that in English constituencies.

Strategic voting in Germany. The (English-language) literature has four main pieces on “ordinary” strategic voting in West Germany – Barnes et al. (1962), Fisher (1973), Jesse (1988), and Bawn (1993) – all employing essentially the same methodology. Each takes the difference between a candidate’s *own* vote total (cast for him or her in a given constituency) and a candidate’s *party’s* vote total (cast for the party list in the same constituency) as a measure of strategic voting.⁹ In particular, a candidate whose own vote falls short of his or her party’s vote is taken to have been strategically deserted. In each case, substantial desertion of small parties is found. For example, Fisher (1973:297–8) reports that 13.5% of the FDP’s list voters deserted the party in the single-member district contests in 1961, with the comparable figure being 29.7% in 1965, and 38.0% in 1969. Jesse’s more extensive study finds FDP desertion rates as high as 61.8% in 1972, 70.9% in 1983, and 61.3% in 1987.

⁸There is also evidence on strategic voting in executive elections. On U.S. presidential elections (not strictly plurality rule but comparable), see e.g. Brody and Page (1973), Abramson et al. 1995. On mayoral elections in Taipei, see Hsieh, Niou, and Paolino (1995). On presidential elections in Mexico and Peru, see Magaloni Kerpel (1994) and Schmidt (1996, N.d.), respectively.

⁹The candidate votes are called *Erststimme* (or “first votes”), the list votes *Zweitstimme* (or “second votes”). Similar analyses also appear in German; see, e.g., Ritter and Niehuss (1987:177–78).

Strategic voting

Casting a list vote for the FDP and a candidate vote for, say, the Christian Democratic Union (CDU) is not unambiguous evidence of “ordinary” strategic voting, however. It may be that the voter truly prefers the CDU, casts a sincere vote for the CDU candidate, but casts her list vote strategically for the FDP, because the FDP is both in alliance with the CDU and in danger of falling below the 5% national threshold (in which case the FDP would get no seats and the CDU might not be able to form a government). There is substantial evidence that supporters of the FDP’s senior coalition partner have deserted their favored party in order to support the FDP (reviewed in Chapter 10). So how is one to tell whether some component of the discrepancy between the FDP’s candidate and list votes is due to “ordinary” strategic voting, intended to avoid wasting the constituency vote?

One approach is to look at surveys that ask voters if they cast split votes and, if so, why. These lend some support to the idea that there is local strategic voting (cf. Roberts 1988:330). Another approach is to look at the district-by-district election returns. If those who give the FDP their list but not their candidate votes are acting for local strategic reasons, then desertion rates should be higher in districts where the contest for the seat is closer. But there is no reason that strategic *list* votes should be cast differentially in constituencies that are close in terms of the candidate votes. Thus, if there is a systematic relationship between the closeness of the constituency race and the FDP desertion rate in each constituency, then this suggests that there are locally strategic voters in Germany too.¹⁰

To investigate this possibility, let the FDP’s percent of the total candidate vote in a given constituency be denoted FDP1, with the FDP’s percent of the total list vote in that same constituency denoted FDP2. Similarly, let GREEN1 and GREEN2 denote the percent of candidate and list votes won by the Greens. The dependent variables in the analyses presented below are two: FDPLOSS = FDP2 – FDP1, measuring the loss the FDP candidate suffers from the baseline set by the party’s list vote; and GRLOSS = GREEN2 – GREEN1, a similar term for the Greens. I regressed each of these dependent variables on MARGIN, the absolute difference between the top two candidates’ vote percentages in the constituency, for the 1987 and 1990 elections. As MARGIN gets larger, the margin of victory in the district gets larger, and the temptation to desert one’s first choice wanes. Thus, a negative coefficient is expected in both cases.¹¹

¹⁰ Another reason often suggested as to why German voters split their votes is that they misunderstand the importance of the *Zweitstimmen*, or “second vote.” There is no reason why this misunderstanding should correlate with the closeness of the constituency race, however.

¹¹ The data for this analysis, along with relevant SAS programs, can be found on the Lijphart Elections Archive’s web site (<http://dodgson.ucsd.edu/lij>) by following the link to “publication-related datasets.”

Strategic voting in single-member single-ballot systems

Table 4.1. Loss of votes by small German parties in constituency contests, relative to list contests

<i>Independent variables</i>	<i>Dependent variable and year of election</i>			
	<i>FDPLOSS 1987</i>	<i>FDPLOSS 1990</i>	<i>GRLOSS 1987</i>	<i>GRLOSS 1990</i>
CONSTANT	5.02 (33.5)	4.36 (27.8)	2.04 (20.3)	-.45 (5.5)
MARGIN	-.04 (5.3)	-.04 (5.1)	-.05 (10.7)	-.02 (4.7)
N =	247	254	247	254
adjusted R ² =	.10	.09	.32	.08

As can be seen in Table 4.1, the coefficient on MARGIN is negative and significant in all four regressions. Not all of this effect is necessarily conventional strategic voting, wherein those who truly rank the third parties first desert them when the district race gets close. Some of it may be due to protest voting by major party voters: If the constituency result is a foregone conclusion, one can take the opportunity to send a pro-environment message to the major parties by voting for the Green candidate.

Strategic voting in Great Britain. The literature on strategic voting in Britain is by far the largest in the world. Much of this literature deals with the elections of the 1980s, when the Alliance surged to near-parity in votes with the Labour Party. Johnston and Pattie (1991) estimate that 5.1% of all voters voted tactically in 1983, with 7.7% doing so in 1987. Heath et al. (1991:54) estimate that “6.5% of major party voters” voted tactically in 1987. Lanoue and Bowler (1992) opine that 5.8% of all voters in 1983, and 6.6% of all voters in 1987, voted tactically. Niemi, Whitten, and Franklin (1992) find these estimates, especially those for 1987, “surprisingly low ... in the light of the efforts of various groups to encourage tactical voting in order to avoid fragmentation of the anti-Thatcher vote.” Interpreting survey responses differently, they estimate that about 17% of all voters were tactical in 1987, a figure which is in accord with an ITN/Harris Exit Poll conducted on election night.¹² Another high-end estimate is offered by Crewe (1987:55), who notes that “among the 23 percent of respondents who claimed to have voted or seriously considered voting Alliance, before deciding against, the

¹²See Evans and Heath (1993) for a critique, and Niemi, Whitten, and Franklin (1993) for a defense of the Niemi, Whitten, and Franklin (1992) methodology.

Strategic voting

overwhelming reason given was some variation of the classic ‘wasted vote’ argument.” Estimates of the percentage of voters that would “consider” voting tactically also vary widely, from an average Gallup figure in 1986–87 of 15% to an average BBC Newsnight figure of 41% (Catt 1989). Even taking the low estimates both of voters that did cast, and voters that would consider casting, a tactical vote, the impact in terms of seats is potentially significant. Butler and Kavanagh (1988:266), for example, reckon that the Conservatives would have won 16 more seats than they did in 1987, had there been no strategic voting.

In addition to estimating the extent of tactical voting, the British literature also explores the determinants of such voting. Lanoue and Bowler (1992) and Niemi, Whitten, and Franklin (1992) both run probit analyses of the probability that individual survey respondents will (report having) cast a tactical vote. Both find that individuals with intense partisan attachments are less likely to vote strategically. This makes sense since intense attachments to one party make it more likely that the other two will be viewed as almost equally bad, which approximates one of the theoretical conditions under which strategic voting is unlikely. Niemi, Whitten, and Franklin also find that respondents whose favorite party was further from contending for the seat, who were better educated, who recalled knowing which party was expected to win, and who had negative feelings about the winning party, were more likely to vote tactically. All these findings fit comfortably with the model of tactical voting expounded above. Voters whose favorite parties ended up out of the running, who were better educated, and who knew before the election who was likely to win in their constituencies, were more likely to know (before the election) that their party was trailing and to have heard the relevant wasted vote argument: hence more likely to have voted strategically. Voters who had negative feelings about the winning party, especially if intense, were more likely to view their second-ranked party as an acceptable vehicle with which to defeat their last-ranked (and clearly threatening) party.

The importance of there being a clear ordering of the second and third candidates is also documented by Galbraith and Rae (1989). Focusing on districts won by the Conservatives in 1983, they find that the swing to Labour (resp. the Alliance) in 1987 was significantly larger if Labour (the Alliance) finished second in 1983. The Alliance swing, for example, was 5.3 percentage points larger on average when the Alliance finished second in 1983 than when Labour did.¹³ Johnston and Pattie (1991)

¹³Galbraith and Rae (1989) find a larger swing to the Alliance despite an artifactual reason to expect a smaller swing. The artifactual reason is this: If the Alliance finished second in 1983, rather than third, then its vote percentage in 1983 was on average larger. A larger 1983 vote percentage, *ceteris paribus*, means a *smaller* swing in 1987.

Strategic voting in single-member single-ballot systems
replicate these findings using a finer-grained measurement of tactical voting and actual vote margins in 1983, rather than just place of finish.

Strategic voting in other countries. In contrast to the plethora of studies of tactical voting in Britain, very little has been written on other countries employing simple plurality, even those whose political conditions approximate those of Britain in the 1980s. There are a few studies of the Canadian and New Zealand experience (cited above), but none that I know concerning India, Trinidad and Tobago, or other developing countries that also use simple plurality.

A brief consideration of the Papua New Guinean experience suffices to show that even simple plurality may not be strong enough to force a sharply divided society into a two-party mold. Papua New Guinea, which became independent of Australia in 1975, has some 700 tribes speaking over 1,000 languages. Its elections, albeit held in single-member districts under simple plurality rules, have not produced any tendencies toward local bipartism. In the 1987 elections, 1,515 candidates chased after 109 seats, with the vote often being fairly evenly divided among the contestants. In the Kerowagi constituency, for example, the winner came in with 7.9% of the poll in a field of 45 candidates. Overall, 41 of the 109 members elected won with less than 20% of the vote (Dorney 1990:57-8). The conditions in Papua New Guinea are almost perfectly designed to discourage strategic voting. With huge fields of candidates, no reliable constituency-level polls, and strong social pressures upon voters to support their own tribes, every candidate (not unreasonably) thinks he may sneak in with a win in a crowded field. Interestingly, what strategic manipulation there is pushes the system toward further fractionalization: "the nomination of 'friendly' candidates to split a powerful opponent's clan vote is a common tactic" (Dorney 1990:59).

Empirical patterns: The bimodality hypothesis

The prediction that third-place candidates will be deserted really holds only in Duvergerian equilibria. 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 ratio of the second to the first loser's vote total – what I shall refer to as the SF ratio. Under Duvergerian equilibria, the SF ratio will be near zero. Under non-Duvergerian equilibria, the SF ratio will be near unity. Thus, if one were to compute the ratio for a number of districts and plot the resulting distribution, one should find a spike at zero and a spike at one.

Strategic voting

Allowing for some frictions in the model – e.g., some noninstrumental voters, some disagreement about which candidates are trailing and which are front-running – the prediction is softened. The SF ratio should either be close to unity (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 zero (when second losers are sufficiently far behind first losers that strategic voting kicks in and they are reduced to their non-instrumental support level, which I assume to be close to zero for most candidates). The SF distribution, in other words, should be bimodal.

A possible real-world example of a non-Duvergerian equilibrium, with an SF ratio near unity, may have occurred in the Ross and Cromarty district of the United Kingdom in its 1970 general election. The final figures were:

Gray	Conservative	6,418
Mackenzie	Liberal	5,617
MacLean	Labour	5,023
Nicholson	Scottish Nationalist	2,268

It is possible, of course, that these figures are the net product of all sorts of strategic calculations by voters – cross-cutting, erroneous, shrewd, etc. Interpreting these results as if they stemmed from a non-Duvergerian equilibrium entails believing the following two points. First, the Liberal and Labour candidates were so close that, before the poll was actually held, it was not at all clear who was in third and who in second; thus, Mackenzie's and MacLean's supporters stuck with them: Neither suffered from strategic desertion. Second, Nicholson did lose his “non-fanatical” support, if any, due to his being obviously out of the running. The 2,264 voters who stuck with him were perhaps those who felt so strongly about the single issue of Scottish independence that they were virtually indifferent between the other three candidates. Alternatively, these voters may have been making an investment in the future, hoping to establish the Scottish Nationalists in their district for a more realistic run at a later time. In either case, they were not short-term instrumentally rational.

I have tested the bimodality hypothesis empirically using data from British elections 1983–1992. Some results, which focus on the behavior of Labour voters, are presented in Figures 4.1 to 4.3. The procedure was as follows. First, I computed the ratio of the vote total of the second loser (third-place candidate) to the vote total of the first loser (second-place candidate) for all districts in which the Conservatives and the Alliance (or its successor, the Liberal Democrats) finished one-two (in some

Strategic voting in single-member single-ballot systems

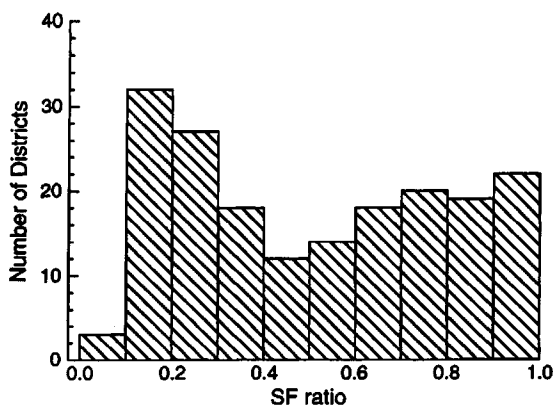


Figure 4.1. Testing the bimodality hypothesis in moderately close districts: British elections, 1983–1992

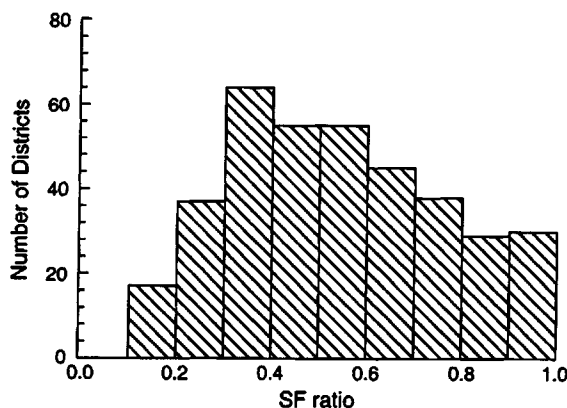


Figure 4.2. Testing the bimodality hypothesis in districts that were not closely contested: British elections, 1983–1992

order), with Labour finishing third.¹⁴ Then I produced a histogram to summarize the distribution of the resulting SF ratios, subject to three different restrictions on the margin of victory in the previous race in the dis-

¹⁴Note that in principle there is a sample selection bias that militates against finding any strategic voting. Only those districts in which third parties decided to field candidates enter the sample. If third parties decide to enter where they think they can hold on to their votes, then the level of strategic voting in the sample will not be representative of the level that would appear were entry decisions exogenous. In practice, this does not appear to be too important, since third parties enter in most U.K. districts.

Strategic voting

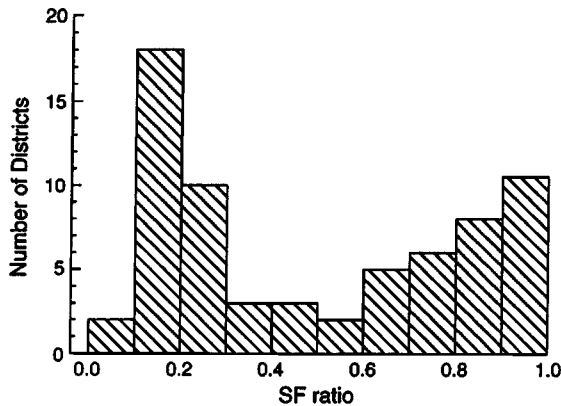


Figure 4.3. Testing the bimodality hypothesis in very close districts: British elections, 1983–1992

trict: that the margin was less than 20% (Figure 4.1), greater than 20% (Figure 4.2), or less than 10% (Figure 4.3).

As can be seen in Figure 4.1, the distribution of SF ratios is bimodal in districts in which Labour ran third and there had been a “close” race (margin less than 20%) the previous time out. In these districts, Labour held on to its support if its candidate was a close enough third, but lost substantial support if its candidate fell too far behind. In contrast, Figure 4.2 shows the SF distribution for districts which, while also featuring a third-place Labour finish, had not been “close” in the previous election. There is no hint of bimodality in the distribution, suggesting that voters do not bother to vote strategically in noncompetitive districts. The importance of a common perception that the race may be close is further suggested by Figure 4.3, which looks just at districts in which the previous race was “very close” (margin less than 10%). As can be seen, the dip in the middle of the SF distribution is even more pronounced in these “very close” districts, suggesting that voters were more willing to vote strategically in more competitive districts. These results comport with previous analyses of the 1983 and 1987 elections (reviewed above) using survey data.

Although the evidence just discussed does indicate there is strategic voting in *some* British constituencies, the constituencies chosen for inclusion in the analysis were those in which it would have made sense for voters to consider a tactical vote (the strategy of investigation here is similar to that in Blais and Nadeau N.d.). If one looks at other districts, one finds much less evidence of strategic voting. Just as the survey evidence shows a distinct minority of the electorate voting strategically – many not being in a position that would logically call for a strategic vote – the aggregate evidence shows a distinct minority of districts with substantial

Strategic voting in single-member single-ballot systems

levels of strategic voting – many not being in a position that would produce larger levels.

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 uncovered in the previous section. 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 front-running candidates, where they are more likely to affect the outcome, rather than to trailing candidates, where they are less likely to affect the outcome. Moreover, allocation or reallocation of resources to front-running candidates should produce the clearest aggregate results (trailing candidates deprived of all instrumental support) when who is trailing and who is not is widely agreed and the margin of victory is small. Thus, the empirical evidence adduced above is far from proving that a significant proportion of the electorate votes strategically. It may be that contributors give only (or mostly) to front-running candidates, or that trailing candidates try to sell their endorsement to front-runners.

The elite-level hypotheses are attractive in that 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 front-running 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 find out who is trailing 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).

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 by-products of everyday activities, such as reading the newspaper, watching TV, 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.

The sensitivity of elite actors to the possibility of strategic voting can be seen in three observations. First, during the 1987 general election in

Strategic voting

the United Kingdom, a group called TV87 formed whose sole purpose was to instruct voters how best to cast a tactical anti-Conservative vote. Their activities consisted primarily in identifying which of the non-Conservative candidates in each constituency were ahead in the latest polls and urging anti-Conservative voters to coordinate on these candidates (Lanoue and Bowler 1992; Catt 1989).

Second, candidates trailing in multicandidate races tend 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. All these actions make good sense from the point of view of preventing their last-place status from becoming common knowledge. These 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 (recall, for example, 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).

Third, candidates who believe they are breaking out of the non-Duvergerian pack into a clear second-place position tend to advertise this fact ostentatiously. Thus, for example, George Bush's crowing about "Big Mo" in the 1980 presidential primaries after his strong finish in the early contests (Bartels 1988) or Merrill Cook's heavy advertising of his second-place poll finishes (as an independent running for the governorship in Utah; see Magleby and Monson 1995).

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 (or attempt to obfuscate this fact). But the voters do the rest: They buy the argument and act accordingly. Empirically, I think that there is substantial evidence that voters have voted strategically in this sense (some of it reviewed above). 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.

4.4 NOMINATION RULES AND STRATEGIC VOTING

Perhaps the clearest example of electoral rules that nullify the alliance-promoting (party-reducing) effect of single-member districts, *even when plurality rule is used*, is encountered in New York state. New York has had a stable multiparty system since the 1940s, despite using plurality rule in single-member districts. The explanation lies in its peculiar mix of rules governing cross-filing, cross-endorsement, and ballot format.

Strategic voting in single-member single-ballot systems

Cross-filing occurs when a candidate for office files not just for his own party's primary but also for one or more others' as well. If a state allows cross-filing, factions within the major parties can open up shop as separate parties without necessarily sacrificing any influence they have in their original party: Their candidates can run in both the minor party's primary and in the major party's primary. California's Progressives took this route in the early part of the century (Scarrow 1986:250).

Cross-endorsement, or fusion, occurs when more than one party nominates the same candidate (and the endorsements appear on the general election ballot). This feature too can provide small parties with an electoral niche to occupy; by regularly nominating one of the major parties' candidates as their own, and stipulating in advance the criteria that will be used in choosing, small parties can influence big parties.¹⁵

The success of this tactic of "auctioning" the small party's endorsement may depend on ballot structure. If a state uses the party-column format, in which all candidates endorsed by a given party appear in a single column with the offices forming the rows, then a cross-endorsed candidate will appear once on the ballot for each party that endorsed him. This allows minor parties to document the size of their voting blocs, since a candidate's total vote will be the sum of his votes in each party's column. In a series of close races, when their support is crucial, this can give small parties considerable bargaining power. If a state uses the office-block format, in which all candidates for a given office appear in a single area of the ballot, together with all their party endorsements, the vote total for the candidate cannot be broken down into subtotals due to each party. The nomination of a small party may still be valuable, but its value is harder to assess.

Since 1947, New York has restricted cross-filing to those who can get the permission of the relevant party's executive committee, allowed unlimited cross-endorsement, and employed a party-column ballot. These three features interact to produce a system quite favorable to the formation and maintenance of minor parties. Small parties can document the size of their vote at general elections and essentially "sell" it (in return for policy or particularistic considerations) to the highest major-party bidder. Thus, what would ordinarily be the left wing of the Democratic party in New York has broken off to form the Liberal party and what would ordinarily be the right wing of the Republican party has broken off to form the Conservative party. Other small parties, of the

¹⁵Currently, ten states allow fusion in state and national elections: Arkansas, Connecticut, Delaware, Idaho, Mississippi, New York, South Carolina, South Dakota, Utah, and Vermont. See Kirschner (1995). I am unaware of any systematic study of the consequences of fusion outside of the New York case, however, except for the important historical studies of Argersinger (e.g., 1980).

Strategic voting

single-issue variety, have found viable niches to occupy as well (Mazmanian 1974; Scarrow 1986).

From the voter's point of view, the New York system can remove any fear of wasting votes by casting them for small parties. So long as a small party supports a viable *candidate*, one also nominated by a major party, their supporters can just as well vote for that candidate under the small party's label as under the major party's label: The candidate's viability is unaffected.

New York shows rather clearly that a single-member plurality system in the general election is no guarantee of "ordinary" bipartism, in which third parties are evanescent and/or politically ineffective. One might argue, then, that the statement of the law needs to be modified to include some explicit conditions on nomination rules; perhaps: "the use of a single exclusive vote in single-member districts operating under plurality rule, together with laws preventing cross-filing and cross-endorsement, tends to produce bipartism." Alternatively, one could simply stress that the logic behind Duverger's Law really does not apply to parties, but rather to candidates, the objects of choice with which voters are directly faced.

4.5 THE ALTERNATIVE VOTE AND MULTIPARTISM

Another way to mitigate the concentrating tendencies of simple plurality rule is to switch from an exclusive to a nonexclusive vote. The effects of such a switch can be seen in elections to Australia's lower house, where the alternative vote (AV) allows a citizen's vote to transfer from one candidate's vote total to another's. The procedure is as follows. As in simple plurality, elections are held in single-member constituencies without secondary districts. Each citizen is required to rank *all* candidates seeking election, from first to last.¹⁶ The returning officer first sorts the ballot papers according to which candidate is ranked first. If at this stage any one candidate has a majority of the votes, he or she is declared elected. Otherwise, the candidate with the fewest first-place preferences is declared defeated. The returning officer then transfers the votes of the defeated candidate's supporters to whichever of the remaining candidates they have marked as their next preference, again checking to see if any candidate has achieved a majority of all votes. This process continues

¹⁶Certain kinds of "mistakes" in ranking candidates are allowed: "A House of Representatives ballot paper is now formal so long as it shows a unique first preference for a candidate and numbers, any numbers, against all the other candidates, or against all the other candidates but one, with the square next to that candidate left blank. Consequently, ballot-papers may be admitted to the scrutiny even when they do not exhibit fully correct numbering, and therefore fail to indicate preferences for all candidates" (McAllister et al. 1990:57).

Strategic voting in single-member single-ballot systems

until some candidate does attain a majority, whereupon he or she is declared elected.

The alternative vote in Australia, like fusion in New York, allows small parties to document their contribution to a larger party's success. It is thus possible, even for parties that virtually never win seats on their own, to play a significant role. Jaensch (1983:21–2) points to three aspects of the Australian electoral system – compulsory attendance at the polls, compulsory ranking of all candidates, and the alternative vote method of translating votes into seats – as underpinning the “blackmail potential” of minor parties. Compulsory attendance at the polls means that minor parties’ potential clientele will turn out and, given that few Australians choose purposely to spoil their ballots, vote. Compulsory ranking of all candidates means that those ranking a minor party’s candidate first will rank *someone* second. This opens the door for the minor party to influence the outcome of the election by issuing “how to vote” cards urging their supporters to adopt a particular ranking of candidates below first. As Jaensch (*ibid.*) puts it, “a minor party which offers (electoral) support in return for (legislative or policy or electoral) concessions, or which threatens electoral retribution if some concession is not offered, must be able to guarantee the allocation of a high proportion of its preferences.” Finally, the AV procedure of counting votes and translating them into seats means that minor party supporters whose party is doomed to elimination at the first round have no reason not to rank their favorite party first. If some party wins on the first count, then they would have done so even had the voter not ranked a hopeless minor party candidate first. If no party wins on the first count, then the voter’s vote will transfer to a more viable candidate.

An example of the viability of very small parties in the Australian system is provided by the Democratic Labor Party (DLP), which flourished 1955–74. Although the party never won a seat in the Australian House of Representatives, “Mackerras (1970) calculated that 81.5 per cent of all DLP second preferences followed the direction of the party and were transferred to the Liberal-Country Party coalition candidates. Further, DLP preferences were instrumental in deciding which party should govern in at least two elections, 1961 and 1969. On both occasions, the coalition government was a ‘second-preference government,’ depending on the DLP” (Jaensch 1983:22).

Despite the hospitality to small parties exhibited by the Australian version of the alternative vote, it would be erroneous to conclude, as is sometimes hinted in the literature, that AV produces no incentives to vote strategically. This conclusion would of course run afoul of the Gibbard-Satterthwaite Theorem’s general guarantee that any democratic voting procedure can generate incentives to vote strategically. Dummett (1984),

Strategic voting

who has considered exactly how strategic voting might arise under AV, points to two main possibilities. First, perhaps one's favorite candidate, while having enough votes to survive the first round, will lose in the second round against one prospective opponent, but probably win against another. In such a case, it behooves one to ensure that the "beatable" opponent is not eliminated in the first round of counting: Thus one may not vote for one's favorite, instead voting for the weaker of two major opponents. Second, perhaps candidate A, who has lots of first preferences and is virtually certain to survive the first round, can defeat your favorite candidate handily but might lose to your second-favorite candidate, who unfortunately has fewer first preferences and is likely therefore to be eliminated in the first round. In this case, it behooves one to vote for one's second-favorite rather than one's favorite.

Note that the first kind of strategic voting under AV, in which one attempts to set up the second round so that one's favorite can win it, does not decrease the effective number of candidates in the first round. Rather just the opposite: One's incentive is to divert votes from a stronger candidate (in terms of first preferences) to a weaker. The second kind of strategic voting, in which one attempts to ensure that a weaker candidate (in terms of first preferences) survives to the second round, may either decrease, leave unchanged, or increase the effective number of candidates. To see this, suppose that one's favorite candidate, C, has 40% of the first preferences, if everyone votes sincerely, while A (whom C cannot beat) also has 40% and B (who can defeat A) has 20%. Suppose also that almost everyone ranking C or A first ranks B second. Depending on whether less than half (but more than one-fourth), exactly half, or more than half of the C-supporters "desert" C and rank B first, the effective number of candidates in the first round will increase, stay the same, or decrease. All three of these cases yield identical *outcomes*: B makes it into the second round, and then defeats C. Thus there is nothing to distinguish them in terms of payoffs. They are all equilibria to the particular game envisioned.

Should we expect strategic voting under AV in practice? On the one hand, voters need more information in order to cast a strategic vote under AV than under ordinary plurality (see Bartholdi and Orlin 1991). On the other hand, some argue that voters will be able to acquire the necessary information and manipulate the system. Dummett (1984:229), discussing the first case above, in which it is necessary to vote for some candidate B in order to prevent another, say C, from surviving the first count, has this to say: "With detailed and reasonably accurate information about the intentions of the voters, such as can be obtained from well-conducted opinion polls, and with a thorough canvass to identify its own supporters, an organized group such as a political party ... can instruct sufficiently many supporters to list A highest to ensure that A is

Strategic voting in single-member single-ballot systems

not eliminated at stage 1, and instruct the rest to list B highest, in order to bring about the elimination of A's principal rival C." He thinks it is "not far-fetched to imagine a political organization's acting in this way," citing the activities of the Birmingham Caucus as a real-world example (albeit under another electoral system). Colin Hughes, co-author of one of the standard references on Australian politics (Hughes and Graham 1968), opines that "tactical voting for partisan purposes is readily understood e.g. when it is advisable to run third so that preferences will be distributed to the less undesirable alternative rather than run second and have that candidate's preferences distributed and go to the more undesirable who would then win" (Hughes 1993:5). I am not aware of any systematic evidence that bears on the frequency of strategic voting of this or other kinds in Australian elections, however.

All told, the case would appear as follows. There is certainly the theoretical opportunity for strategic voting under AV, and there is some expert opinion that it appears in practice. But more information is needed to vote strategically under AV than under simple plurality. And, whereas strategic voting always acts to decrease the effective number of candidates under simple plurality, it is as likely to increase as to decrease this figure under AV. Thus, small parties can be viable under AV where they would not be under simple plurality. AV does not exert as strong a reductive influence on the party system as does simple plurality.

4.6 CONCLUSION

This chapter has investigated strategic voting in single-seat elections held under a variety of single-ballot procedures. In the process, I have sought to specify the theoretical and institutional conditions under which Duverger's Law does and does not hold at the local level. My conclusions can be summarized as follows:

Institutional limits on Duverger's Law. In the last three sections, I have considered the U.S. system (a single, exclusive, non-fused candidate vote; cast in a single-member district without secondary districts; decided by plurality rule; with fusion candidacies outlawed), the New York system (identical to the U.S. system except that fusion candidacies are allowed), and the Australian system (identical to the U.S. system except that the vote is nonexclusive and a majority is required for election). It is not usual to describe an electoral system by listing such a long train of features. But each item in the list is arguably necessary to produce local bipartism. Approval voting differs only in that there are multiple votes; and many believe that it would lead to multipartism, although there is no empirical evidence on this score. The Australian alternative vote sys-

Strategic voting

tem differs only in that the vote is nonexclusive (and in the use of majority, rather than plurality, rule); and Australia has more than two significant parties. The SNTV system used formerly in Japan differs only in that the districts were multimember rather than single-member; and postwar Japan has had a multiparty system except briefly in the mid-fifties. The system used in Germany in 1949 differs only in that there were secondary districts; and Germany at that time had a multiparty system. It is hard to imagine changing the electoral formula holding all else constant. Ignoring differences in the voting options available to voters, however, one might say that the French system differs only in that runoff elections are held if no candidate garners an absolute majority of votes, rather than awarding the seat to the candidate with a relative majority (i.e., a plurality); and France has a multiparty system. Finally, New York's system differs only in that fusion candidacies are allowed; and New York has a multiparty system.

One possible lesson of this exercise is that Duverger's Law really pertains to a quite specific system and is not very robust to small changes in that system. Another possible lesson is that there are many ways one might improve the prospects of smaller parties, and hence promote multipartism. Some ways (increasing the primary district magnitude, or adding a secondary district) entail also improving the overall proportionality of the system. Some ways (making the vote nonexclusive, introducing runoffs, allowing fusion candidacies) do not. Finally, one might conclude that the importance of the plurality formula in promoting bipartism has been exaggerated. It is obviously not a sufficient condition for bipartism (witness New York or West Germany 1949). Nor, in light of Austria, Malta, Colombia, and Uruguay – all of which have had long spells of two-partyism, despite having one form or another of PR – is plurality a necessary condition.

Theoretical limits on Duverger's Law. Suppose that one focuses on the ordinary plurality system originally considered by Duverger. Does the logic of strategic voting play out at the local level as he suggested? Many in the literature take this for granted, convinced by the usual wasted vote argument. In this chapter, I have specified the preconditions that must be met for strategic voting to have much impact and also noted that strategic voting need not necessarily appear in equilibrium in three-candidate races.

Consider the behavioral preconditions of the model first. The model shows that the extent to which strategic voting winnows out weak candidates depends on how many short-term 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.

Strategic voting in single-member single-ballot systems

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 losing. If clear information about candidate chances is provided, one can expect substantial levels of strategic voting and a consequent reduction in the number of viable candidacies. If little (or conflicting) information is provided, then greater amounts of sincere voting (or cross-cutting strategic voting) can be expected, and the tendency toward two viable candidates will be weaker.

Consider next some preconditions of the model concerning the structure of partisan preferences and competition. One such precondition stipulates that not too many voters can have a clear first choice but be essentially indifferent between the rest of the field (since such voters have no incentive to vote strategically). Another precondition, first noted by Riker (1976), forbids the existence of a party that is a sure winner. As one example of when sure winners might arise, consider a polity in which the structure of political competition is really unidimensional, and the largest single party stands athwart the median position in most constituencies. In this case, leftist and rightist voters are "stuck." Even if a supporter of a leftist party notes that her party is out of the running, supporting a larger rightist party does not further her interests and supporting a larger leftist party will still leave that party in, at best, second place (either because the centrist party has enough votes on its own to defeat a coalition of parties on the left, or because, if it does not, it will attract sufficient right-wing support to defeat the leftist challenge). Thus, voters facing such a structure of competition might as well vote sincerely.

Even if all the preconditions of the model are met, the result that follows is still a bit more hedged than the typical formulation in the literature. It is true that the most likely equilibrium in the pure model is a Duvergerian one, in which third parties are devastated by strategic voting. But non-Duvergerian equilibria can arise when two or more candidates are tied for second, because in this case neither will be obviously "out of the running," and hence their supporters will have no clear incentives to desert them. In the pure model these non-Duvergerian equilibria arise only with precise ties for second, and appear to be generally unstable (Fey 1995). But if voters perceive larger variances in candidates' vote shares than they do in the pure model, then near ties (where what counts as "near" is defined relative to the perceived variance in candidate vote totals) may suffice to forestall any clear shaking out of the field of candidates. The present

Strategic voting

model thus provides specific and empirically testable predictions about what kind of exceptions to local bipartism one should expect – something that has not previously been done in a systematic fashion.

Empirical evidence for local bipartism. Most of the evidence adduced in the literature relating to Duverger's Law has been national-level and, as noted in Chapter 2, the evidence is not very supportive. What is pertinent in assessing the local bipartism argument, of course, is evidence at the constituency level. Here it is certainly possible to find contrary evidence: In recent general elections in Britain and New Zealand, for example, three or more candidates have received significant vote shares in a number of districts. But it is also possible to find ample positive evidence of strategic voting playing the vote-concentrating role attributed to it in the standard view, when key conditions are met. Thus, although there are clear theoretical conditions that limit the force of the local bipartism argument, these conditions can be and are met or approximated sufficiently often to make strategic voting an important force, pushing party systems toward bipartism as Duverger argued.

There are of course other possible avenues to explore in explaining local bipartism. Duverger appropriately suggested 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 trailing to front-running candidates is 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).

Beyond bipartism. Finally, I should note that the wasted vote argument does not imply local bipartism, as Duverger and others in the literature have asserted. The argument does provide a reason to expect downward pressure on the number of competitors, in case there are more than two, as shown in this chapter. But, although I have spoken here of the “local bipartism” result, the wasted vote argument does not in fact provide any reason to expect upward pressure on the number of competitors, in case there is only one. I elaborate on this point in the next chapter.