POLITICAL ECONOMY OF INSTITUTIONS AND DECISIONS

Editors
James E. Alt, *Harvard University*
Douglass C. North, *Washington University of St. Louis*

Other books in the series
Alberto Alesina and Howard Rosenthal, *Partisan Politics, Divided Government and the Economy*
Lee J. Alston, Thrainn Eggertsson, and Douglass C. North, *Empirical Studies in Institutional Change*
James E. Alt and Kenneth Shepsle, eds., *Perspectives on Positive Political Economy*
Jeffrey S. Banks and Eric A. Hanushek, *Modern Political Economy: Old Topics, New Directions*
Yoram Barzel, *Economic Analysis of Property Rights, 2nd Edition*
Robert Bates, *Beyond the Miracle of the Market: The Political Economy of Agrarian Development in Kenya*
Peter Cowhey and Mathew McCubbins, *Structure and Policy in Japan and the United States*
Gary W. Cox, *The Efficient Secret: The Cabinet and the Development of Political Parties in Victorian England*
Jean Ensminger, *Making a Market: The Institutional Transformation of an African Society*
Murray Horn, *The Political Economy of Public Administration: Institutional Choice in the Public Sector*
Jack Knight, *Institutions and Social Conflict*
Michael Laver and Kenneth Shepsle, *Making and Breaking Governments*
Michael Laver and Kenneth Shepsle, *Cabinet Ministers and Parliamentary Government*
Brian Levy and Pablo T. Spiller, *Regulations, Institutions, and Commitment*
Leif Lewin, *Ideology and Strategy: A Century of Swedish Politics (English Edition)*
Gary Libecap, *Contracting for Property Rights*
Mathew D. McCubbins and Terry Sullivan, eds., *Congress: Structure and Policy*
Gary J. Miller, *Managerial Dilemmas: The Political Economy of Hierarchy*
Douglass C. North, *Institutions, Institutional Change, and Economic Performance*
Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*
J. Mark Ramseyer, *Odd Markets in Japanese History*
J. Mark Ramseyer and Frances Rosenbluth, *The Politics of Oligarchy: Institutional Choice in Imperial Japan*
Jean-Laurent Rosenthal, *The Fruits of Revolution: Property Rights, Litigation, and French Agriculture*
Charles Stewart III, *Budget Reform Politics: The Design of the Appropriations Process in the House of Representatives, 1865–1921*
George Tsebelis and Jeannette Money, *Bicameralism*
John Waterbury, *Exposed to Innumerable Delusions: Public Enterprise and State Power in Egypt, India, Mexico, and Turkey*
David L. Weimer, *The Political Economy of Property Rights*

MAKING VOTES COUNT

STRATEGIC COORDINATION IN THE WORLD'S ELECTORAL SYSTEMS

GARY W. COX
University of California at San Diego

CAMBRIDGE
UNIVERSITY PRESS

1

Introduction

Early in the 1984 presidential primary season in the United States, it was clear that the sitting President, Ronald Reagan, would easily win the Republican nomination and that former Vice President Walter F. Mondale was the front-runner for the Democratic nod. Democratic voters who knew that they disliked Mondale faced a coordination problem: If all of them could agree on a single alternative to Mondale, from among the half-dozen or so candidates languishing in single digits in the opinion polls, they could conceivably deny Mondale the nomination; but if they failed to agree on a single alternative, then Mondale would almost surely win. Although anti-Mondale Democrats shared a dislike of Mondale, they differed substantially in their preferred alternative. Thus, even putting aside the complexities of the American primary process, it was by no means clear *ex ante* that anti-Mondale Democrats could coordinate on an alternative. In the event, although Gary Hart emerged as the focal alternative to Mondale and enjoyed a large and rapid run-up in the polls, his candidacy faltered and Mondale secured the nomination.

Early in the 1990 presidential campaign in Peru, it was clear that Nobel Prize-winning novelist Mario Vargas Llosa was the front-runner. Peruvian voters who knew that they disliked Vargas Llosa faced a coordination problem: If all of them could agree on a single alternative to Vargas Llosa from among the half-dozen or so candidates trailing in the polls, they could conceivably deny Vargas Llosa the presidency; but if they failed to agree on a single alternative, then Vargas Llosa would almost surely win. Although anti-Vargas Llosa voters shared a dislike of Vargas Llosa, they differed substantially in their preferred alternative. Thus, it was by no means clear *ex ante* that anti-Vargas Llosa Peruvians could coordinate on an alternative. In the event, Alberto Fujimori rocketed from obscurity late in the campaign to become the focal anti-Vargas Llosa candidate, securing a strong second-place finish in the first round of voting, then defeating Vargas Llosa in the runoff (Schmidt N.d.).

These two examples illustrate several general features of electoral coordination: the mixture of common and opposed interests; the possibility of success or failure; and the rapidity with which vote intentions change when coordination takes off. The examples' focus on strategic voting in presidential elections is too limited, however. Modern representative democracy presents at its core a *series* of coordination problems that arise as natural consequences of electoral competition for governmental offices. A group with enough votes to elect some number of candidates in a given (legislative or executive) race will in fact elect that number only if it can make its votes count by concentrating them appropriately. One way to avoid spreading votes too thinly is to limit the number of candidates. But which potential candidates, representing what shades of opinion, will withdraw in favor of which others? If attempts to limit the number of candidates fail, another chance to make votes count arises on polling day, when voters can concentrate their votes on a subset of the available candidates. But which candidates will bear the brunt of strategic voting and which will be its beneficiaries?

This is a book about strategic coordination broadly conceived, covering both legislative and executive elections, both strategic entry and strategic voting. It investigates the consequences of strategic coordination and those structural features that determine the nature of the coordination problems that political actors face in differing polities.

The consequences of strategic coordination. Successful electoral coordination reduces the number of electoral competitors. When leftist elites agree to join together into a single leftist party, rather than continuing with some larger number, there are fewer parties nominating fewer legislative candidates. If leftist elites do not coordinate their endorsements sufficiently, leftist voters may complete the coalition that the elites tried but failed to form, by deserting one of the leftist candidates for the other(s). In the process they decrease the effective or vote-weighted number of candidates.[1] Duverger's famous Law – the proposition that "the plurality rule [employed in single-member districts] tends to produce a two-party system" (Duverger 1954:113) – is a claim about how far the processes of reduction can be expected to go in the case of one particular set of electoral rules.

Electoral coordination is not just a matter of reducing the number of parties competing in elections, however, any more than coordination on

[1]Just as an industry with 100 firms, one of which makes 95% of all sales, is essentially a monopoly despite its 100 firms, so one might say that an election with 100 candidates, one of whom garners 95% of the vote, has not much more than one "real" candidate. The notion of an "effective number of parties," due to Laakso and Taagepera (1979), is one attempt to count "real" candidates. If v_j is the vote share of the jth party, then the effective number of parties is $(\sum v_j^2)^{-1}$, the reciprocal of the Hirschman-Herfindahl concentration index.

technical standards is just a matter of reducing the number of such standards. When writers of software programs agree on standards compatible with Microsoft's operating system, this does reduce the sales-weighted number of operating systems, and may even lead to the withdrawal of some operating systems from the market. But, in addition, there are some winners (Microsoft; those who like PCs) and some losers (Apple; those who like Macintoshes). Similarly, when leftist opinion leaders agree to rally around Socialist Party A's candidates, rather than around Socialist Party B's, this does reduce the vote-weighted number of parties, and it may even lead to the disappearance of B from political competition. But, in addition to any gain of seats that the unified socialists may accrue as a whole, there are some relative winners (party A; those who prefer its policies) and losers (party B; those who prefer its policies). To put the point more starkly: Successful electoral coordination necessarily involves a reduction in the number of competitors; but such a reduction just as necessarily entails a selection of which competitors will survive, and this selection potentially has important policy effects.

In this book, I shall consider both the reductive and the redistributive effects of electoral coordination. The reductive effect of strategic coordination is most evident when it succeeds, the redistributive effect most evident when it fails – as will be seen.

The nature of the electoral coordination problem. As regards what determines the nature of the coordination problem that arises in any given system, I shall be principally concerned with three main independent variables: electoral institutions, political motivations, and public expectations. The importance of the first of these factors – electoral institutions – has been alternately asserted and dismissed since Duverger's seminal work in the 1950s (Duverger 1954). Here, electoral institutions – which determine the available opportunities for trading votes in order to win more seats – are taken as largely defining the coordination game that elites and voters must play.

Electoral institutions are not the whole story, however. A second part of the strategic situation is defined by the preferences of the elite and mass actors who must coordinate. If leftists care mostly about policy, and hate each other's policies almost as much as they hate the current government's, then there is little incentive for them to coordinate their actions, even if by so doing they could win more seats. If leftists care substantially about future elections, then it may be a good strategy to play tough in the early rounds, enduring a series of coordination failures in the hopes of emerging eventually as *the* leftist party.

Finally, expectations are crucial in any game of coordination, and electoral coordination is no different. If Socialist Party A believes that B's

supporters will vote strategically (for A), in the event that both enter, then A has little incentive to acquiesce in any demands that B might make. If B has opposite beliefs, there is no room for the elites to resolve the coordination problem on the Left. As for the voters, if poll results clearly reveal that A's candidates are ahead, then B's supporters will more likely desert to A than the reverse. If polls are absent, noncredible, or ambiguous, however, then the informational prerequisites of strategic voting may not be satisfied, in which case one again expects a failure of coordination.

Of the three independent variables just mentioned, the first – the nature of the electoral institutions in a polity – is obviously central to comparative electoral studies. The second – the nature of political actors' preferences – is a standard concern, especially of rational choice scholars. The third variable, however, concerning the nature of actors' expectations, may be less obviously relevant to some readers. Before proceeding further, let me say something more about the role that expectations play in elections.

1.1 ELECTORAL SYSTEMS AS SYSTEMS OF EXCHANGE: THE ROLE OF EXPECTATIONS

It is conventional, but no less compelling for that, to express wonder at the vast array of activities that are coordinated by the market and its attendant price system. Somehow, without any central planner dictating that it be so, about the right amount of food descends on New York City, about the right number of flashlights make their way to Omaha, and about the right number of video cassette recorders arrive at Gila Bend.

The key to the process by which consumer demands are anticipated and fulfilled with such enviable accuracy (at least by central planning standards) is the system of prices. Clearly known prices for intermediate and final goods and services allow a vast decentralization of planning and productive activities. *Market-clearing* prices, attained in the hypothetical equilibria of economic models, equate demand and supply. At those prices, the number of widgets that consumers in the aggregate seek to purchase turns out to equal the number of widgets that businesses in the aggregate seek to sell (ignoring inventories and other subtleties).

Political scientists do not usually think of elections as systems of exchange subject to equilibrating mechanisms. But there are some analogies between the exchange of voting support among citizens within the electoral system and the exchange of consumer goods among citizens within the market. Relative to the imaginable extreme in which everyone runs for president and votes for him- or herself, real-world presidential

elections are highly concentrated and coordinated affairs. In the United States, everyone expects that only a handful of Republican or Democratic politicians are viable candidates for their parties' nominations, and they act accordingly. Contributors do not contribute to, activists do not volunteer for, and citizens do not vote for hopeless candidates, ensuring that those expected to do poorly, do poorly in fact. Somehow lots of people, with diverse preferences, are willing to contribute in various ways to Bob Dole's candidacy but not to Pete Wilson's. Bob Dole accordingly is willing to continue as a candidate; Pete Wilson is not.

The key to the process by which voter demands are anticipated and fulfilled is the system of expectations. Clearly known common expectations about who is and is not viable are self-fulfilling, and allow a considerable decentralization of planning and vote-productive activities. *Market-clearing* expectations, attained in the hypothetical equilibria of political models, equate demand and supply. At those expectations, the number and type of candidates that voters are willing to vote for turns out to equal the number and type of candidates that are willing and able to stand for election.

Equilibrium, whether economic or political, may of course be a rare bird. Too many entrepreneurs may set up fast-food restaurants in a given (geographical) location, leading to poor (expected or realized) profits and a shake-out in the industry. Too many politicians may set up candidacies at a given (ideological) location, leading to poor (expected or realized) vote totals and a contraction in the field. Developers anticipating a large influx of population may play Chicken against one another in building housing tracts to fulfill the anticipated demand.[2] Groups anticipating a large anticommunist vote in a post-communist eastern European election may play Chicken against one another in launching campaigns to attract the anticipated votes. All of these examples illustrate dynamic adjustment on the supply side, or what happens to the supply of goods or candidacies when prices or expectations are not sufficiently clear.

One could also adduce examples of demand-side informational failures. Consumers are unaware of a spiffy new product that is cheaper and better than a well-advertised alternative that everyone currently uses; it takes some time before word of mouth moves market demand toward the new product. Leftist voters are unclear as to which of two leftist candi-

[2]The original game of Chicken pits two teenagers in hotrods against one another. Both head down the center of the road toward each other, the first to swerve being "chicken." If neither swerves, a very bad outcome results. If one swerves, then the swerver is humiliated while the other is covered with glory. If both swerve, an intermediate payoff results.

dates is ahead (or behind) in a three-way race also including a right-wing candidate. As a result, the leftist vote is split and the rightist wins the seat.

Such evident and important nonequilibrium examples notwithstanding, equilibrium analysis has been fundamental to market economics for some two hundred years. One of the premises of this book is that equilibrium analysis ought also to be fundamental to the understanding of elections. Although there is some work that fits the broad description of equilibrium analysis outlined above, in which expectations play a central role in coordinating electoral activity and choice, there is no book-length treatment of the subject that attempts to explain how different electoral laws affect the nature of market-clearing expectations and electoral coordination. The present work seeks to begin filling that gap.[3]

1.2 PLAN OF THE BOOK

A brief outline of the book can now be given in terms of the three independent variables introduced above – electoral institutions, political motivations, and public expectations. Electoral institutions determine how votes translate into seats. If political actors care mostly about winning seats in the current election, then the influence of electoral institutions on their goals is direct. If, in addition, actors' expectations about each other's vote shares are precise and consensual, then a well-structured coordination game emerges in which the prospects for successful coordination are good. This model corresponds to the standard Duvergerian approach to legislative elections in the electoral studies literature.

Parts II and III of the book formally generalize the Duvergerian model of strategic coordination – both at the level of citizens coordinating votes and elites coordinating endorsements and entry – to legislative electoral systems other than the single-member simple plurality case for which the logic is best developed in the extant literature. Using a formal model forces one to state assumptions explicitly. This leads almost immediately to fairly substantial changes in the way that one understands even so well-known a result as Duverger's Law. For example, although Duverger and the subsequent literature have been quite clear in saying that plurality rule leads to bipartism, the only valid conclusion from the arguments they explicitly advance is that the number of viable parties cannot exceed two. *More generally, in any electoral system the necessity of electoral coordination only implies an upper bound on the number of competitors.*

[3]Perhaps the clearest examples of the kind of equilibrium analysis suggested above are the complete information models of Osborne and Silvinski (1995), Besley and Coate (1995), and Feddersen (1992). In this book, I shall focus more on incomplete information models.

Recognizing this simple fact leads to a number of changes in the way that one thinks about the impact of electoral laws on party systems. For example, I argue in Chapters 7 and 10 that the correct understanding of the institutionalist model implies that the number of parties in a system ought to be an interactive function of electoral and social structure. Many have viewed Duvergerian institutional analysis as reading social cleavages out of the analysis (a point upon which I expand in Chapter 2). But a closer look at what the institutional analysis really entails reads them back in.

Beyond clarifying the general nature of the impact that electoral institutions have – they impose upper bounds, rather than pushing systems toward some specific equilibrium number of parties – this book also identifies what the appropriate upper bounds are. Part II derives the bounds imposed by strategic voting in a range of different electoral systems: single-member simple plurality (SMSP), single nontransferable vote (SNTV), proportional representation (PR), and others. Part III brings strategic entry into view, again focusing on the restraining effect of the electoral system.

Part IV turns to aspects of electoral coordination that hinge on the executive choice procedure, again assuming for the most part that agents are interested primarily in winning office in the current election and that expectations are consistent. Duverger argued that the desire of voters in single-member simple plurality elections to avoid wasting their votes meant only that there would be pressure toward local bipartism in each legislative district. He had an additional argument, developed later by Sartori (1968; 1976), as to why a congeries of potentially unrelated local bipartisms might cumulate into national bipartism. Part IV deals with this systemic part of the institutionalist argument – putting the stress not on the formation of national parties (as did Duverger and Sartori) but instead on competition for executive office.

Part V turns away from the model in which agents are assumed to care primarily about winning seats in the current election and to have aligned expectations about who is best positioned to do so. When other motivations and expectations are entertained – agents that care about current and future policy outcomes, rather than just current seats, for example – the probability of coordination failure increases. The most obvious consequence of coordination failure is not so much that the number, or effective number, of competitors goes up (as it does) but that whichever side of the political spectrum has failed more egregiously to coordinate pays a penalty in seats. If the Left splits in a single-member district, the Right wins the seat. In Part V, I investigate coordination failures and how they affect the quality of representation, the maintenance of dominant parties, and the politics of realignment.

Having briefly sketched the sequence of topics to be dealt with, let me next say something about methodology. This study differs from previous works in comparative electoral studies both in its reliance on formal game theoretic analysis of the incentives set in train by different electoral institutions and in its use of primarily district-level data to test the hypotheses that the theory entails. In Section 1.3, I discuss the use of formal theory in electoral studies, with particular reference to strategic voting. Then, in Section 1.4, I comment briefly on the data that I use in this book.

1.3 THEORIES, FORMAL THEORIES, AND ELECTORAL STUDIES

The study of mass voting systems has been carried on in two distinct theoretical traditions. One tradition, originally a part of mathematical economics and philosophy, can itself be broken down into work in *social choice theory* (e.g., Arrow 1951; Sen 1970; Fishburn 1973; Gibbard 1973; Schwartz 1986), *public choice theory* (e.g., Buchanan and Tullock 1962; Mueller 1989), and *spatial theory* (e.g., Downs 1957; Hinich, Davis, and Ordeshook 1970; Romer and Rosenthal 1979; Palfrey 1984; Cox 1990a; Enelow and Hinich 1990) – to mention only some of the better-known categories. This work uses the tools of formal symbolic logic, mathematical welfare economics, microeconomics, or game theory to get where it is going. A second tradition, the domain of political scientists and sociologists, is characterized by the work of such scholars as Duverger (1954), Rae (1971), Sartori (1976), Lijphart (1984, 1994), and Taagepera and Shugart (1989). It is less formal, more engaged with real-world data, and more interested in concrete political problems – while still being theoretical for all that.

These two traditions seldom speak to one another, as Dummett (1984) and Reeve and Ware (1992) have observed. Research into strategic voting provides an illustrative case of parallel (nonintersecting) development.

Within the electoral studies tradition, concern with strategic voting arose because it was believed to reduce the number of political parties competing in some systems. Duverger's original formulation (1954) seemed to be that strategic voting was *present* in simple plurality systems, acting to push them toward bipartism, whereas it was *absent* in PR and majority runoff elections, which in part explained their tendency toward multipartism. Reacting to Duverger's apparent belief that his "psychological factor" was inoperative under PR systems, Leys (1959:139) and Sartori (1968:278) argued that strategic voting under PR was no different in kind from that found under plurality, differing only in the degree to which it came into play – and, hence, in the degree to which it tended to reduce the number of viable parties in the system.

Sartori's notion of a continuum of systems, from *strong* (in which strategic voting and elite coalitional activity act forcefully to depress the number of parties) to *weak* (in which strategic voting and incentives to form coalitions are largely absent and thus put little downward pressure on the number of competitors), is now standard in the literature.

Within the formal theoretic tradition, concern with strategic voting was sparked by Arrow's theorem, which presumed that social choice processes could operate on the true preferences of the citizenry. The work of Gibbard (1973) and Satterthwaite (1975) demonstrated formally that incentives to vote strategically could arise in any minimally democratic voting system, and the Gibbard-Satterthwaite theorem has since become a benchmark result in the literature.[4]

It is clear that the Leys-Sartori conjecture (they offered no proof of their assertions) and the Gibbard-Satterthwaite theorem are similar: Both assert the general existence of strategic voting incentives across a wide range of voting systems. Nonetheless, neither side cites the other. Gibbard and Satterthwaite were undoubtedly completely unaware of the Leys-Sartori conjecture. No other formal theorists have since recognized Leys and Sartori as precursors. Returning the compliment, one can read the post-Gibbard/Satterthwaite classics of electoral theory, even those which give substantial attention to strategic voting – such as Taagepera and Shugart (1989) or Lijphart (1994) – without finding any mention of Messrs. Gibbard and Satterthwaite.

One might say that this does not matter. After all, if one asks whether Taagepera and Shugart (or Lijphart) exhibit some fundamental flaw in their approach to strategic voting, due to their not using formal theory in their books, the answer is that they do not. If one asks whether Gibbard and Satterthwaite suffered from not having the kind of detailed knowledge upon which Leys and Sartori based their assertions, the answer is that they did not.

Nonetheless, as I am peddling formal theory in this book, and also using the insights of the electoral theory tradition, let me say something about what each has to offer. If one compares the Gibbard-Satterthwaite theorem to the Leys-Sartori conjecture, the theorem wins hands down in terms of rigor and precision. But it is not as useful to political scientists as it might be, because its conclusion is politically ambiguous. The theorem merely alerts one to the possibility that there may be strategic voting under any democratic electoral system, while saying nothing about either the political consequences of that strategic voting, or about how much strategic voting one should expect. In contrast, the Leys-Sartori

[4] I have stated the result loosely. A careful discussion of the Gibbard-Satterthwaite theorem at an elementary level can be found in Ordeshook (1986:82-86).

conjecture focuses on a particular kind of politically relevant strategic voting – the kind that acts to reduce the vote-weighted number of parties – and says something specific about which systems will have a lot and which a little. This greater relevance presumably explains why political scientists who study electoral systems are more likely to use Sartori's distinction between strong and weak systems than they are to cite the Gibbard-Satterthwaite theorem.[5]

In this book, I hope there will be a fruitful combination of traditions. My *interest* is largely in the questions raised by the electoral theorists; my *methods* are largely those of the formal theorist. Thus, although the formal models that I shall use look at strategic voting broadly conceived, including both strategic voting that does and does not depress the number of parties, it is on the former kind that I focus. Moreover, in each model I seek to say something about the equilibrium level of strategic voting. The result, in that part of the book dealing with strategic voting, is a series of formal theorems each of which looks like a version of either Duverger's Law or the Leys-Sartori conjecture, restricted to a specified range of electoral systems.

1.4 DISTRICT-LEVEL DATA

Another gap that the present work seeks to begin filling is that between our electoral theories (mostly district-level) and data (mostly national-level). As Taagepera and Shugart (1989:117) note, "most studies of electoral systems ... have dealt with the whole system rather than with the district level." This book departs from that tradition, in that most of the data employed are district-level rather than national.

A substantial impediment to conducting electoral research with district-level data is, of course, finding the data in a machine-readable form. As part of the research for this work, I have directed an effort to expand and computerize the Lijphart Elections Archive at my home institution, the University of California at San Diego. The result of that effort, along with most of the data used in this book, can be found on the World Wide Web at http://dodgson.ucsd.edu/lij.

[5]Formal theorists have not entirely ignored the issues in which electoral theorists are interested. There is, for example, a large formal literature that investigates how much strategic voting one should expect under different systems. One approach, due to Nurmi (1987), ranks voting systems in terms of the amount of information about preferences that a voter needs in order to cast an intelligent strategic vote. Under simple plurality, one needs to know the vote intentions of the other voters. Under a majority runoff system, one needs to know a bit more: others' vote intentions both in the first and in the second round. Under the single transferable vote system, one needs to know yet more: Bartholdi and Orlin (1991) show that it is "NP-complete" to determine how to vote strategically; colloquially, this means that it is horrendously difficult except in a few special cases.

2

Duverger's propositions

Students of politics have asked how electoral laws affect the formation and survival of political parties since mass elections first became common in the late nineteenth and early twentieth century. Henry Droop, an English advocate of proportional representation (and inventor of the Droop quota), noted as early as 1869 that plurality elections promote what later scholars have called "strategic" or "tactical" voting:[1]

> 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 other of the parties between whom the election really lies (quoted in Riker 1982:756).

Droop was also surely aware of the plurality system's tendency to underrepresent minority parties – a topic generally discussed today under the rubric of "disproportionality," "big-party bias," or the "mechanical effect." In any event, by 1881 he had enunciated a version of what is now called Duverger's Law:

> the only explanation which seems to me to account for [the two-party systems in the United States, United Kingdom, etc.] is that the two opposing parties into which we find politicians divided in each of these countries have been formed and are kept together by majority [what we now call plurality] voting (quoted in Riker 1982:756–7).

Duverger himself originally thought in terms of three propositions, one for each of the main electoral systems in use when he wrote (Duverger 1986:70). Eventually, however, he settled on just two, which

[1]Duverger referred to the same phenomenon under the rubric of the "psychological factor," by which he meant the desire of voters to avoid casting a "wasted vote" for a candidate with no hope of winning.

Riker (1982) has dubbed Duverger's Law and Duverger's Hypothesis. Duverger's Law states that "the simple-majority single-ballot system [i.e., simple plurality rule] favors the two-party system" (Duverger 1954:217). Duverger's Hypothesis states that "the simple-majority system with second ballot and proportional representation favors multipartyism" (Duverger 1954:239).

The rest of this chapter considers Duverger's propositions at length, proceeding as follows. Section 2.1 reviews some well-known criticisms of Duverger's work advanced by those who take a sociological approach to the study of party systems. Since much of the rest of the book is devoted to studying the political consequences of electoral laws, it is important first to address the arguments of those who doubt that these consequences are particularly important. Section 2.2 then clarifies the particular version of Duverger's propositions upon which the first part of the book will focus – a version framed at the district level in terms of the effective number of candidates, rather than at the national level in terms of the actual number of parties. Section 2.3 reviews the standard logic underlying Duverger's propositions, having to do with strategic voting decisions in the mass electorate, on the one hand, and strategic coalition decisions in the elite strata, on the other. Section 2.4 offers a preliminary sketch of when and how strategic voting appears in systems other than simple plurality, with particular attention to how this might place an upper bound on the number of viable candidates. Section 2.5 concludes with an outline of Part II of the book.

2.1 SOCIOLOGICAL CRITIQUES OF DUVERGER'S PROPOSITIONS

Scholarly reaction to Duverger's work has been highly polarized. Two particularly sharp disagreements – one over his research's scientific status, another over its causal validity – illustrate this polarization clearly.

As regards the scientific status of Duverger's propositions, opinion could not be more divided. On the one hand, some question whether Duverger's generalizations serve "any useful function at all" (Jesse 1990:62; cf. Wildavsky 1959:318) or dismiss them on fundamental grounds: the impossibility of summing up complex and reciprocal social interactions in scientific laws (Lavau 1953; Mackenzie 1957; Bogdanor 1983:261). Sartori (1994:30) views this camp as predominant, noting that "the prevailing wisdom of the profession still is ... that comparatively valid generalizations are impossible to achieve." On the other hand, Riker has devoted an entire essay to the thesis that research into Duverger's Law exemplifies scientific progress (Riker 1982). And Sartori

(1994:27), while disagreeing with Riker's specific formulations, asserts that the arguments of those skeptical of the possibility of comparative generalization "are demonstrably wrong."

Scholarly opinion regarding the causal validity of Duverger's propositions is similarly divided. Two main controversies have arisen. First, some argue that Duverger simply mistook the direction of causality. In this view, party systems determine electoral systems, rather than the other way around (Grumm 1958; Eckstein 1963:253; Lipson 1964; Särlvik 1983:123; Fukui 1988:121). The central body of evidence supporting this view has been contested by Duverger and Riker, but a comparison of Riker's (1982) and Bogdanor's (1983) views of the *same* evidence reveals the two sides far from agreeing.

A second (and closely related) challenge to the causal validity of Duverger's work holds that he focused on an unimportant variable. In this view, party systems are determined primarily by the number and type of cleavages in society, with electoral structure playing either an inconsequential, or at least a distinctly secondary and variable, role (Campbell 1958:30-32; Grumm 1958; Lipson 1959; Meisel 1963; Lipson 1964; Lipset and Rokkan 1967; Rokkan 1970; Blondel 1972:237; Nohlen 1981; Beyme 1985; Franco 1986:82–3; Solari 1986:120–21). Lavau (1953:46), for example, in perhaps the earliest retort of this kind to Duverger's theses, opined that "the method of voting remains a rather small consideration among the complex and infinitely diverse factors that, combined differently in each national society ... condition political life." In the purer forms of this school of thought, long-term multiparty systems such as those found in Europe are to be explained by the existence of many strong social cleavages – cleavages which would find expression in the party system even under single-member plurality rules. Conversely, long-term two-party systems such as that in the United States are to be explained either by inherent social dualism (e.g., Charlesworth 1948; Key 1964b:229ff) or by the relative mildness of their social and ideological cleavages (e.g., Lipson 1953; Hartz 1955). Set against the "social determinist" school, especially its purer variants, are a number of scholars who, while admitting (to varying degrees) the importance of social structure, still view electoral structure as having a consistently important independent impact (e.g., Riker 1982; Duverger 1986:71; Taagepera and Shugart 1989:53; Sartori 1994).

The reason that Duverger's Law has stuck in the craw of so many political scientists of a sociological bent is that it seems to set up some sort of "institutional determinism," wherein markedly different social cleavage structures are hypothetically all mashed into one final outcome (a "two-party system") merely upon application of a particular set of

electoral laws. Social cleavages thus seem to play no *systematic* role in determining the equilibrium number of parties. They do play a residual role – Duverger states his law as a tendency precisely to allow for the possibility that particularly strong social cleavages might retard the reduction in number of parties that single-member plurality systems promote. But this puts the cart before the horse in the eyes of those who see social cleavages as exogenous and strongly determinative, electoral laws as endogenous and (at best) marginally determinative.

It is no less true that the stronger versions of "social determinism" stick in the craw of institutionalists, especially those with roots in economic rather than sociological theory. A belief that socially defined groups will always be able to organize in the political arena seems to ignore the problem of collective action (Olson 1965), and a belief that they will always organize *as parties* seems to say that "going it alone" is always a better strategy than forging coalitions. Moreover, the number of social cleavages seems large relative to the number of parties in *any* society, so how is one to tell which cleavages are big enough to be party-defining and which are not? Is it obvious, for example, that the cleavage between Finnish and Swedish people in Finland (which gives rise to a separate party) is more intense than that between European-Americans and African-Americans in the United States (which does not)? From an institutionalist perspective, politicians can take socially defined groups and combine or recombine them in many ways for political purposes (Schattschneider 1960). A given set of social cleavages does not imply a unique set of politically activated cleavages, and hence does not imply a unique party system.

Duverger's propositions are obviously controversial and the institutional determinist and social determinist positions are obviously far apart. Yet, despite the wide divergence of views articulated in the literature, the battle between proponents (institutionalists of various stripes) and detractors (mostly political sociologists) has been anything but sustained and focused. In part this has to do with the very size of the disagreement. To the extent that one side of a debate (e.g., Riker 1982) takes social "science" as clearly possible and desirable, while at least some on the other side (e.g., Lavau 1953; Mackenzie 1957) take it as problematic and perhaps not even desirable, it is difficult to find common ground.

I shall not have much to say here about whether scientific study of comparative politics is possible. I certainly hope that it is, but a convincing defense of this hope against more pessimistic views is well beyond the scope of the present work.

One need not take the argument over Duverger's Law all the way to first principles in the philosophy of science, however. The sociological

critiques carry weight even if one is committed to a scientific approach to comparative politics. Accordingly, I address these criticisms – concerning the endogeneity of electoral structure and its relative unimportance when compared to social cleavages – before proceeding further. I shall in particular be interested in the degree to which the two views can be reconciled or synthesized.

The endogeneity of electoral structure

Some reconciliation between institutionalist and sociological perspectives is certainly possible in the controversy over whether electoral systems cause party systems or vice versa. Indeed, there is a close symbiotic relationship between these two claims. On the one hand, if electoral laws do indeed affect the ability of political parties to survive as independent organizations, as Duverger's propositions imply, then presumably parties will seek to manipulate those laws to their own advantage when they can. Assuming Duverger is right, in other words, leads naturally to the conclusion that the party system (and the calculations of partisan benefit rattling around within it) may affect the electoral system. On the other hand, the claim that parties tinker with the electoral mechanism in order to ensure their survival, or increase their vote totals, presupposes a belief on their part in electoral engineering. There would be no point in seeking a new electoral system if electoral systems did not matter. Thus, the early sociological attack, if it is to hang together logically, must take for granted that electoral laws confer partisan advantages, or at least that parties believe they do, in order to conclude that parties will attempt to change them.

It is true that the endogeneity of electoral laws can take some of the causal steam out of electoral structure. For, if electoral systems can be changed relatively easily, then one might expect frequent changes for short-term partisan gain, as in Greece, France, or Turkey. Frequent changes – or the anticipation of change – might then undercut the long-term causal effects of electoral law. Consider, for example, a small party facing a single-member plurality system. If such a party believes the electoral system will survive unaltered for a long time, it faces substantial incentives to join or form a coalition capable of securing a plurality. If, on the other hand, the party believes the electoral system can be easily changed, it may simply seek to bring about such change. The strength of the incentive to coalesce produced by an electoral system depends, in other words, on how long that system is expected to last.

This is a good point. It may be what the sociological critics had in mind. But it is not the death knell of the causal validity of Duverger's Law.

I think the main lesson to be learned from the endogeneity of electoral structure is that in some systems, such as Greece and the emerging democracies of Eastern Europe, the perceived changeability of the electoral code may dampen the incentive effects of electoral law. There is some reason to think these cases may not "fit" quite as well with the older, more established electoral systems. Perhaps one could even get some mileage out of these cases in measuring how much perceived changeability affects the reductive strength of a system.[2]

But in fact electoral laws are not everywhere and always easily changed. It is the winners under the current electoral system who (if party discipline holds) must find it in their interest to change the electoral system.[3] Thus – unless there is substantial uncertainty against which even the winners wish to insure themselves;[4] or the winners think the electoral situation has changed, so that the old electoral rules will no longer serve them well, and they can agree on how to manipulate the rules for short-term gain (e.g., France in 1951 and 1986); or the electoral system has come to symbolize an unpopular political regime, so that politicians face intense public pressure to rewrite the code (e.g., the recent reforms in Japan, Italy, and Venezuela) – electoral systems tend to be, and to be perceived as, rather long-lived. As Lijphart (1994:52) puts it, "one of the best-known generalizations about electoral systems is that they tend to be very stable and to resist change." When electoral systems *are* (perceived as) hard-to-change and long-lived, however, the incentives they set in train are no longer discounted by the probability that the rules will change.

[2]Even in cases where the electoral code is seen as protean, there may be ways to decide whether electoral structure pushes the party system or vice versa. The problem is a garden-variety one of simultaneous or reciprocal causation. It does indeed make it impossible to infer from a simple bivariate correlation something solid about causation. But there are in principle solutions to such problems. One is to fashion a simultaneous equations model, in which both the electoral system and the number of parties appear as endogenous variables. No one in the literature has attempted to do this, and I do not propose to do so here; but it is not obviously an impossible route to take, merely a very difficult one! Were one to take this route, it would be important to find some variables that predict the adoption of electoral systems that are not endogenous to social structure. In this regard, the findings of Blais and Massicotte (N.d.) are interesting: In a study of the determinants of electoral law in some 166 countries, two of the strongest predictors of a country having a single-member plurality system that they find are "being a former British colony" and "size in km^2." The first is arguably exogenous to indigenous social structure. The second may be a weak proxy for social diversity but in that case the correlation runs in the opposite direction from that which would be predicted from a social determinist perspective.

[3]In some cases, electoral laws can be changed by executive decree (as in Russia 1993) or popular initiative (as in the recent repeal of the senate electoral law in Italy), in which case the "winners" may not control the process of electoral change.

[4]I am thinking here of the uncertainty that European politicians faced upon the introduction of universal suffrage near the turn of the century, an uncertainty that seems to have played an important role in the introduction of PR. Cf. Carstairs (1980); Noiret (1990).

Social cleavages and the party system

Consider next the importance of social structure. As Taagepera and Grofman (1985:343) note, "it is much harder to find testable propositions" regarding social structure in the literature. Nonetheless, they believe that some in the literature can be taken "as standing for the proposition that 'the more axes of cleavage there are within a society, the greater will be the number of political parties.'" Nohlen (1993:27), for example, offers a thesis in which the number of social cleavages affects not just the number of political parties but also (following Grumm 1958) the nature of the electoral system a given country will possess:

> the greater the social fragmentation, the more probable is the adoption of a proportional [electoral] system and also the rise of a multiparty system. The greater the social homogeneity, the more probable is the adoption of the simple plurality system; but, also, the more probable is the rise of a two-party system ... or of a limited party pluralism.

Although a bit fuzzy, the idea that social cleavages condition the party system has considerable force and has spawned an entire literature in opposition to, or at least in tension with, the institutionalist literature. The prospects for a limited reconciliation of the institutionalist and sociological perspectives are, however, reasonably good. Two points in particular are worth noting at the outset.

First, to assert that social structure matters to the formation and competition of parties – which no one denies, when the point is stated in such a broad fashion – does not imply that electoral structures do not matter. To make this latter point, one has to adopt a rather extreme monocausalist perspective according to which the underlying cleavage structure of a society is so much more important than the details of electoral law that basically the same party system would arise regardless of the electoral system employed (cf. Cairns 1968:78). Does anyone believe that the United States would remain a two-party system, even if it adopted the Israeli electoral system?

Second, to assert that electoral structure affects party competition in important and systematic ways does not imply that social structure is irrelevant. It might appear that this is exactly what Duverger's Law does imply – bipartism in any society merely upon application of single-member districts – but in fact that overstates Duverger's proposition and the institu-

tionalist development of it, where there has been an increasing appreciation of the interaction effects between social and electoral structure.

In the next two subsections, I expand on the two points just made. I first offer some systematic evidence that electoral structure matters, even controlling for social structure. Then I discuss how social and electoral structures might interact in the process of party formation and maintenance.

The importance of electoral institutions. In this section, I compare the number of parties competing (and winning) in elections to the upper and lower houses of those countries that possess elective upper chambers. If social cleavages drive the number of parties competing, with little strategic adjustment to electoral rules, then one should find essentially the same number of parties competing for votes in both house and senate elections, regardless of the electoral systems used in the two bodies. If, on the other hand, groups do adapt to the electoral environment, then there should be predictable differences in the number and effective number of competitors.

Comparing the number of parties in house and senate elections would seem to be a natural method of controlling for social diversity, as the society in which the elections are held is the same. Nonetheless, as far as I know no one has bothered actually to compute the numbers and make the comparisons. I do so by identifying all democratic countries that possessed elective upper chambers circa 1990, analyzing their electoral systems to arrive at an *a priori* expectation regarding the number of parties that should compete in house and senate elections, and then testing these expectations against the empirical record.

Circa 1990 there were 16 democratic countries that possessed elective upper chambers, consisting of 7 Latin American, 5 European, 2 Anglophone, and 2 Asian cases. In Table 2.1, I list these countries, along with a brief characterization of the electoral systems used in each house.[5] The middle column in the table indicates which system should have the larger number (or effective number) of parties. For example, in Australia the house and senate are both elected under single transferable vote (STV) rules but the senate districts all return more than one member, while the house districts are all single-member. Thus, one expects a greater number of parties competing in the Australian senate than in the Australian house.

Before proceeding, I should note four caveats about the predictions listed in Table 2.1. First, Uruguay is too complicated to yield a clear prediction – at least I have not been able to come to a clear *a priori* expectation. Second, the predictions I do make are merely ordinal: They state which chamber should have more parties, without stating how large a

[5]Fuller descriptions of the lower house electoral systems in most of these countries are given in Chapter 3.

difference in number there should be. In some cases, the difference would seem to be fairly large (e.g., Australia) while in others it would seem to be quite small (e.g., Belgium). Third, four of the countries listed employ a fused vote for legislative races. That is, voters have only one vote to cast for a slate that includes both house and senate (and sometimes presidential) candidates. For this reason, the prediction for these systems is necessarily different when it comes to votes as opposed to seats. The parties competing for house and senate seats, and the votes they receive, must be equal in fused vote systems; it is only in the translation of votes into seats that the two chambers can differ. Fourth, one would hardly expect that the party systems for house and senate elections would fully adapt to their respective electoral systems, in splendid isolation from one another. If a party can run and elect candidates under the more permissive system, it may decide to run candidates in the other system as well – not to win seats, perhaps, but to keep its electoral organization in good trim, to establish its blackmail potential, or for other reasons. In this case, the party system in each chamber should be influenced by that of the other, in such a way as to lessen observed differences.

Bearing these caveats in mind, Table 2.2 displays the (effective) number of parties winning seats. As can be seen, of the 15 countries with clear predictions, 14 (93%) show the expected difference in the effective number of parliamentary parties,[6] while 11 (73%) show the expected difference in the scalar number of parliamentary parties.[7] For countries with fused votes, the differences are purely mechanical effects. For the other countries, the difference presumably understates the mechanical effect due to strategic adaptation.

Table 2.3 displays the (effective) number of parties winning votes. Of the 12 countries with predictions, 11 (92%) show the expected difference in the effective number of elective parties, while 10 (83%) show the expected difference in the scalar number of elective parties. If one removes the fused vote systems from the analysis, since they must trivially agree with the prediction of equal numbers, there are 8 countries with predictions, of which 7 (88%) and 6 (75%), respectively, are in conformity with the prediction.

As noted above, not all of the comparisons contrived in Tables 2.2 and 2.3 are such that one expects a large difference in the number of parties; sometimes the electoral systems for house and senate are very simi-

[6]The only exception, the United States, is trivial in that the electoral systems for the house and senate are identical, the only difference (and the reason for the prediction) being that the house has a larger number of districts than the senate.

[7]The exceptions in terms of scalar number of parties are Belgium, Colombia, Poland, and Spain. Belgium is not much of an exception in that the house and senate have very similar electoral systems and are equal in the number of parties competing.

Table 2.1. Comparing the electoral rules for house and senate elections in sixteen countries, circa 1990

Country	Electoral rules, house	Prediction	Electoral rules, senate
Australia	STV with $M = 1$	<	STV with $M > 1$
Belgium	See Chapter 3.	>	Same as lower house system, with lower district magnitudes. Indirectly elected members excluded.
Bolivia[a]	PR with median magnitude = 13.	>	List plurality with $M = 3$.
Brazil	Open list PR with median magnitude = 11.	>	Plurality rule in 1- and 2-seat districts
Chile	Open list PR with $M = 2$. 60 electoral districts.	>	Open list PR with M = 2. 19 electoral districts.
Colombia	PR with median magnitude = 6.	<	PR with $M = 100$.
Dom. Republic[a]	PR with median magnitude = 2.	>	Plurality rule with $M = 1$.
Italy	See Chapter 3.	>	Similar to house system, with smaller district magnitudes.
Japan	SNTV with $M = 3$, 4, or 5.	<	76 seats elected by SNTV with M between 1 and 4. 50 seats elected by PR with $M = 50$.
Philippines	Plurality rule with $M = 1$.	>	Nationwide plurality election of 12 senators. Each voter has 12 votes.
Poland	PR with median magnitude = 10.	>	Mostly 2-seat districts, voters having two votes each (non-cumulative), and the top two vote-getting candidates winning the seats.
Spain	PR with median magnitude = 5.	>	4-seat districts in which each voter casts 3 votes. Some indirectly elected members.
Switzerland	PR with median magnitude = 6.	>	Plurality rule with $M = 1$ or 2.

Table 2.1. (cont.)

Country	Electoral rules, house	Prediction	Electoral rules, senate
U.S.	Plurality rule with $M = 1$.	>	Plurality rule with $M = 1$.
Uruguay[a]	See Chapter 3.	?	PR with $M = 30$.
Venezuela[a]	PR with district magnitudes varying according to population; up to 5 additional seats awarded on the basis of national vote totals.	>	Same system with lower district magnitudes ($M = 2$), fewer additional seats (3), and fewer total members to be elected.

[a]These countries have fused votes.

lar. This is the case, for example, in Belgium. Thus, although Belgium is technically an exception in terms both of the effective and the scalar number of elective parties, it is not much of an exception. On the other hand, not much difference should be expected in the United States or Chile, either, and so the "successes" there ought to be somewhat discounted too. On the whole, the pattern of evidence is consistent with the notion that different electoral systems do produce different party systems, even when used in the same society at the same time.

The interaction of social and electoral structure. Duverger took social structure more or less as a residual error, something that might perturb a party system away from its central tendency defined by electoral law. Later scholars, however, have considered the possibility that cleavage and electoral structures may interact. For example, two recent papers that take this tack – Kim and Ohn (1992) and Ordeshook and Shvetsova (1994) – both come to the conclusion that Duverger's institutionalist claims are conditioned by the nature of social cleavages.[8]

Kim and Ohn elaborate a point made previously by Sartori (1968), Rae (1971), and Riker (1982) in order to accommodate the Canadian

[8]Another paper that plies the same waters is Taagepera and Grofman (1985). They argue that Duverger's propositions work only if there is one dominant social cleavage, and even then they offer some emendations. The cleavages about which they talk, however, are really *politicized* cleavages, not all cleavages in the society, whether brought into political significance or not. Cf. Ordeshook and Shvetsova (1994:107).

Table 2.2. Comparing the number and effective number of parties winning seats in the house and senate of sixteen countries, circa 1990

Country	ENPP – house	NPP – house	Prediction	ENPP – senate	NPP – senate	Year of house/ senate election
Australia	2.03	3	<	2.57	5	1993
Belgium	8.28	13	>	8.24	13	1991
Bolivia[a]	3.92	5	>	3.43	4	1989
Brazil	8.6	19	>	5.5	6	1990
Chile	5.06	10	>	4.68	7	1989
Colombia	2.18	6	<	2.22	5	1990
Dom. Republic[a]	3.06	4	>	2.23	3	1990
Italy	5.60	10	>	3.88	5	1992
Japan	2.46	7	<	3.66	8	1990/ 1989
Philippines	3.46	7	>	2.42	5	1992
Poland	3.88	7	>	3.56	12	1993
Spain	2.67	11	>	2.58	14	1989
Switzerland	6.52	14	>	3.44	7	1987
U.S.	1.94	3	>	1.96	2	1992
Uruguay[a]	3.30	4	?	3.24	4	1989
Venezuela[a]	4.65	8	>	3.98	5	1988

Main Sources: Central Intelligence Agency 1994; Nohlen 1993; Mackie and Rose 1991. Different sources handle independent and minor party candidates or lists differently (e.g., one might group them all in a single 'other' category, another might separate 'independents' from 'minor parties'). These differences mean that one cannot use the data in the table to make confident comparisons between countries. I have tried, however, to ensure that the data for the house and senate from a given country are handled comparably, to ensure that that comparison is meaningful.

[a]These countries have fused votes.

exception to Duverger's Law (Canada has simple plurality elections yet a long-standing multiparty system). They point out that one of the suppositions underlying Duverger's Law – that small parties will be underrepresented under plurality rule in single-member districts – depends for its validity on the geographic distribution of voters. In particular, if a third party's supporters are concentrated in a particular region of the country, then they may be able to compete successfully as one of the two main parties locally, even while remaining a third party nationally. Thus, Duverger's Law holds only if the social cleavage structure is not characterized by geographically concentrated minorities who might form the basis of a successful, albeit localized, third party.

Table 2.3. Comparing the number and effective number of parties winning votes in the house and senate races of sixteen countries, circa 1990

Country	ENPV – house	NPV – house	Prediction	ENPV – senate	NPV – senate	Year of house/ senate election
Australia	2.47	4	<	2.61	6	1993
Belgium	9.75	13	>	9.79	13	1991
Bolivia[a]	5.01	10	=	5.01	10	1989
Brazil	9.7	34	>			1990
Chile	7.22	19	>	5.43	17	1989
Colombia	2.22	8	<	2.26	7	1990
Dom. Republic[a]	3.92	4	=	3.92	4	1990
Italy	6.18	10	>	4.10	5	1992
Japan	2.91	8	<	4.47	10	1990/1989
Philippines	3.32	5	>	2.05	4	1992
Poland			>			1993
Spain	4.37	9	>			1989
Switzerland	6.80	18	>			1987
U.S.	2.07	3	>	1.99	2	1992
Uruguay[a]	3.37	5	=	3.37	5	1989
Venezuela[a]	3.36	9	=	3.36	9	1988

Main Sources: See Table 2.2.

[a]These countries have fused votes and therefore the numbers for the house and senate are identical.

Ordeshook and Shvetsova (1994) reanalyze Lijphart's (1990) data with an eye to clarifying how social structure matters in determining the number of parties. They find that the number of parties in a country increases with the diversity of the social structure and with the proportionality of the electoral structure, but also that these effects interact. Increasing the proportionality of an electoral system in a homogeneous society does not proliferate parties, whereas it does in heterogeneous societies. Similarly, increasing the diversity of the social structure in a non-proportional electoral system does not proliferate parties, whereas it does in a proportional system.

If institutionalists have sometimes explored the importance of social cleavages, it is no less true that those with primary interests in political

sociology have also recognized some "institutional" constraints on the formation of parties. Meisel (1974) and Jaensch (1983, ch. 3), for example, are at pains to point out that not all social cleavages become politicized, and that even fewer become *particized* (i.e., made into important lines of partisan division). Both processes – politicization and particization – typically do not just happen; they require someone to push them along, someone with resources who can compete against other political entrepreneurs who may be attempting to prevent the politicization of that particular cleavage, or to activate others instead.

Taking both social cleavage structures and electoral structures into account, there are three key stages to consider when accounting for the level of vote or seat concentration observable in any particular polity. The first stage is the translation of social cleavages (here taken to be exogenous but obviously susceptible to political manipulation[9]) into partisan preferences. The second stage is the translation of partisan preferences into votes. The third stage is the translation of votes into seats.

In some institutionalist models, the first stage is not explored: There is an exogenously given number of parties with clear demarcating features (e.g., the position they adopt along an ideological dimension), so that voters' preferences over parties are easily deducible. No party ever fails to get votes because it is too poor to advertise its position; no would-be party ever fails to materialize because it does not have the organizational substrate (e.g., labor unions, churches) needed to launch a mass party. In an expanded view, of course, the creation of parties and the advertisement of their positions would be key points at which a reduction of the number of political players occurs. The multiplicity of possible or imaginable parties is reduced to an actual number of *launched* parties, then to a smaller number of *known* parties, even before the electorate produces an effective number of vote-getting parties, and the electoral mechanism produces an effective number of seat-winning parties.

(1) The reduction of possible to launched parties depends on many things: the level of preexisting nonpolitical organization that can be turned to political advantage; monetary resources; media access; and so on. Thus, a religious cleavage with well-organized and well-financed churches on both sides (e.g., Evangelicals versus Pietists in the nineteenth-century United States) is more likely to be politically activated, other things equal, than a racial cleavage in which one side is poorly organized and poorly financed (e.g., whites versus Aborigines in Australia).

[9]Ethnic and linguistic identities can be manipulated. Laitin (1994), for example, gives an example of how British colonial policy gave tribal chieftans in Ghana an incentive to accentuate their linguistic differences.

(2) The reduction of launched parties to known parties depends primarily on media access and money. It is unlikely that Screaming Lord Such's party in England would have done much better had it been even more widely known than it was but in principle lack of exposure is a stumbling block for many minor parties. Whether this stumbling block can be overcome may depend partly on strategic decisions by potential contributors: If all of them seek to avoid wasting their contributions on hopeless (because unknown) parties, then the party will remain unknown (because poor), hence helpless.

(3) The reduction of known parties to voted-for parties is the domain of strategic voting. Even if known, a party still has to be viable in order to attract votes.

(4) Finally, the reduction of voted-for parties to seat-winning parties is typically a mechanical feature of the electoral system. The only substantial exceptions occur in systems in which votes are not pooled across all candidates from a given party, as in Taiwan or Colombia. In these systems, the distribution of a party's vote support across its candidates or lists materially affects its seat allocation (Cox and Shugart N.d.).

It is obvious that social cleavages matter. Institutionalists now have a small start toward specifying what kinds of social cleavages matter, and how, under different kinds of electoral systems. If this line of research is continued within the institutionalist paradigm, then institutionalists and political sociologists may have more to say to one another in future. In any event, I think that the interaction between electoral systems and social cleavages merits further research.

2.2 NARROWING THE FOCUS

Having convinced the reader, hopefully, that the sociological critiques of Duverger's propositions do not compel their abandonment, the next step is to clarify the version of these propositions upon which I shall focus in the first part of the book. There are by now a good many versions of Duverger's propositions in the literature and they differ consequentially along at least two dimensions: whether the dependent variable is defined at the national/system level or at the district level; and whether the dependent variable concerns entry deterrence, post-entry winnowing of the field of candidates, or both. I shall discuss these two distinctions in turn.

The national versus the district level

For Duverger, the dependent variable in his Law and Hypothesis was the number of "serious" parties (somehow defined) at the national level. Most political sociologists also take a national view, a fact which may

explain why many of them are underwhelmed by the evidence for Duverger's Law. For, if one looks only at the national level for evidence, one finds relatively few examples of electoral systems employing single-member districts and plurality rule. Ignoring small states (below 5,000,000 population), the list would as of 1992 include just seven democratic polities from Chapter 3's list of 77: Bangladesh, Canada, Nepal, New Zealand, the United Kingdom, the United States, and Zambia. To this list, one might add the Philippines and India (which fail to meet the criteria in 1992 but have substantial democratic experience). Of these nine, only the United States has recently had reasonably pure two-party electoral competition. Zambia is dominated by one party. The rest have multiparty electoral competition, and in most the "third" parties succeed in securing a significant presence in the legislature. Thus, by this accounting, there is no empirical regularity to explain.

In contrast to those who concentrate on the national level, there are also those who shift the focus of Duverger's Law to the district level. Duverger himself wrote that "the true effect of the simple-majority system is limited to *local* bi-partism," that is, "the creation of a two-party system inside *the individual constituency;* but the parties opposed may be different in different areas of the country" (Duverger 1955:223; italics added). Nonetheless, he attempted from that beginning to extend the argument to the national level, and of course stated his sociological laws at that level. Leys (1959) and Wildavsky (1959) were perhaps the first to question the validity of Duverger's extension, and to insist that his propositions operated only (in the case of Leys) or at best (in the case of Wildavsky) at the district level.

From a district-level perspective, Duverger's Law is supported every time one finds a district that is dominated by two parties. One can disaggregate even further and look not at a string of elections in a given district over time, and the pattern of party competition that characterizes that period, but instead at individual district elections. At this level of disaggregation (geographically to a single district, temporally to a single election), the relevant prediction is that the top two parties will together garner most of the votes in the election. The district-level evidence is not entirely unproblematic but nonetheless looks a good deal more impressive than does the national-level evidence.

In the first part of this book, I take a completely disaggregated view of electoral systems and their effects. That is, I focus on the processes whereby most district races end up with a limited number of viable candidates (or lists), rather than on the processes whereby polities end up with a limited number of viable parties. Connections between these two levels of analysis (e.g., how the strategies of national parties may perturb

district results from what they might be were districts somehow insulated from national forces, how district-level incentives percolate up to the national level, etc.) are considered later.

Pre-entry versus post-entry politics

In addition to focusing on the district rather than the nation, I shall also focus on post-entry rather than pre-entry politics. In the post-entry period, after some number of candidates has entered the fray, the central question concerns the processes that winnow the field out – strategic voting, strategic contributing, and so forth. In the pre-entry period, before candidacies have been announced, the central question concerns the processes that deter entry – the major parties' nomination procedures (which facilitate coordination on a single nominee), the anticipation of failure by third-party and independent candidates, and so forth. Another way to put the distinction is that in the pre-entry period one watches as an indefinitely large field of potential candidates is reduced to a definite field of actual candidates, while in the post-entry period one watches as an actual number of entering candidates is reduced to a smaller effective number of vote-getting candidates.

The effective number of candidates, due to Laakso and Taagepera (1979), is the reciprocal of the Hirschman-Herfindahl index widely used in the industrial organization literature to measure how concentrated sales in a given industry are. It is now the standard measure of how concentrated vote shares are in electoral contests.[10] If there are K actual candidates in a race, the maximum possible effective number of candidates is also K (a value attained when all K candidates garner a 1/K share of the votes). As votes concentrate on a smaller number of candidates, the effective number of parties falls below the actual number. If votes concentrate to the point that only two candidates get any, an extreme version of Duverger's prediction, then the effective number of parties will be bounded above by 2. If voters merely tend to concentrate on two candidates, then the effective number may be somewhat above 2.

2.3 THEORETICAL EXPLANATIONS OF DUVERGER'S PROPOSITIONS

Given our theoretical focus – upon single elections in particular districts, after the field of candidates has been established – the pertinent question

[10]If v_i is the vote share of the ith candidate, the effective number of candidates is $1/\Sigma v_i^2$.

for Duverger's Law is as follows. Why do voters in one-seat districts (in which each has a single vote to cast, the candidate with the most votes winning) tend to concentrate their suffrages on only two candidates?

One answer is strategic voting. Instrumentally rational voters eschew wasting their votes on hopeless candidacies, preferring instead to transfer their support to some candidate with a serious chance of winning. As long as everybody agrees on which are the hopeless candidacies, strategic voting will mean that votes concentrate on the serious candidates, of whom there will usually be just two in equilibrium (see Chapter 4).

Another whole set of answers can be generated by noting that *any* class of agents (not just voters but also opinion leaders, contributors, party officials, etc.) will tend to allocate whatever resources they control (not just votes but also endorsements, money, campaign appearances, etc.) to serious rather than to hopeless candidates. All that one needs for this conclusion is that the agents in question be instrumentally rational – that is, motivated primarily by a desire to affect the outcome. For then contributions to hopeless candidacies will be pointless, as they are unlikely to affect the outcome.[11]

Scholars disagree over which of these two causal mechanisms – strategic voting in the mass electorate or strategic contributing in the elite strata – is the more important. On the one hand, some argue that strategic voting is irrational, given the infinitesimal chance that a single vote will affect the outcome, and conclude that most of the action must be at the elite level (see Meehl 1977; Riker 1982:764). On the other hand, there is considerable evidence that voters do behave strategically, and at least one study (Gunther 1989) finds elites bungling their strategic role.

In my view, both kinds of resource concentration are important. Elites typically act first: Contributions and endorsements are sought before votes are. If elites coordinate fully, on just two candidates, then the voters are left with a binary choice and, accordingly, vote sincerely. If the elite strata fail to coordinate fully, then a multicandidate field will typically be further winnowed by strategic voting within the electorate (typically instigated by the prospective elite beneficiaries).

2.4 STRATEGIC VOTING AND THE NUMBER OF VIABLE COMPETITORS

Long ago Leys (1959:139) and Sartori (1968:278) argued that strategic voting should appear even under PR systems, to the extent that those sys-

[11]It is true that contributions large enough to single-handedly convert hopeless into serious candidacies may still be instrumentally rational; but few agents control such resources.

tems embodied significant departures from pure proportionality (due to small district magnitudes, high thresholds, or large premiums awarded to major parties). Leys put the matter as follows: "Although most writers seem to assume that [strategic voting] has no place in the analysis of any PR system, a weakened [form of strategic voting] may be highly plausible" (p. 139). Sartori advanced a very similar thesis, noting that "the influence of PR [through its encouragement of strategic voting] merely represents an enfeeblement of the same influence that is exerted by the plurality systems." Both authors thus placed single-member simple plurality systems and the various real-world PR systems on a continuum as regards their tendency to reduce the number of viable political parties below the theoretical benchmark number that would flourish under a purely proportional system (see also Rae 1971:141–43).

Part II of this book generalizes Leys' and Sartori's claim. Vote concentration due to the strategic diversion of resources away from hopeless candidates is not unique to single-member districts operating under Anglo-American rules. *Any* electoral system can be characterized by an equilibrium upper bound on the number of candidates (or party lists), such that if the actual number exceeds this upper bound there is a tendency for instrumentally rational voters to concentrate on a smaller number. In the remainder of this section, I shall informally sketch out the steps through which one must go in order to identify the equilibrium upper bound of candidates (lists) implied by strategic voting. Following chapters formalize the argument for a certain number of electoral systems.

The first step in identifying the equilibrium number of candidates (lists) for a particular electoral system is to ask how many candidates (or lists) one would expect to be seriously in the running for a seat (or seats). Given instrumentally rational voters, if a candidate (list) is *not* seriously in the running, then he (it) will lose support. Thus, whenever the number of candidates (lists) exceeds the number that can plausibly be seriously in the running for a seat, one expects some of the candidates (lists) to suffer from strategic desertion by their followers.

So how does one decide how many candidates (lists) can be "seriously in the running" in a given electoral system? The most obvious answer regarding candidates running in M-seat districts under plurality rule (top M finishers get seats) is $M + 1$. This answer is defended at length in Chapter 5, but to get a preliminary idea of why $M + 1$ might be the right answer, consider a race between $K > M$ candidates for $M = 5$ seats. Suppose first that the rules of election are those once employed in Japan: Each voter casts a single nontransferable vote, the five candidates with the most votes winning seats. There are two ways that a voter in such a system can "waste" her vote: She can vote for an almost-sure loser; or

she can vote for an almost-sure winner. Neither vote is likely to improve the outcome from what it would otherwise have been had the voter abstained. Thus, instrumentally rational voters will avoid both kinds of waste, sending their votes to the "marginal" candidates – those who have, relative to nonmarginal candidates, a large probability of being tied for *M*th, hence of being on the edge between winning and losing.

Imagine now that a random sample of voters is polled and it is found that the proportion of the electorate intending to vote for candidate j is π_j, where candidate labels are chosen so that $\pi_j \geq \pi_{j+1}$ for all $j < K$. If $\pi_{M+1} > \pi_{M+2}$, and the gap between $M + 1$ and $M + 2$ is large relative to the sampling error in the poll, then $M + 2$'s probability of being tied for *M*th is infinitesimal, relative to $M + 1$'s. Voters can quite confidently count $M + 2$ out (as, for that matter, can contributors, party leaders, and so forth). Any resources controlled by instrumental agents will therefore flow away from $M + 2$ and toward the marginal candidates (which set includes, in equilibrium, the top $M + 1$ candidates). Thus, to the extent that the gap between the first loser ($M + 1$) and the second loser ($M + 2$) is "noticeable," one expects voters (and other contributors) to concentrate their resources on the strongest $M + 1$ candidates.

Now suppose that the rules are changed to largest remainders with the Hagenbach-Bischoff quota ($Q = V/(M + 1)$, where V is the total number of votes cast for all lists). How many lists will be serious contenders? We can again identify the *candidates* in order of their chances at winning a seat and being marginal (tied for *M*th). For example, the lead candidate on a list for which the expected vote share exceeds the quota is likely to win, and becomes virtually certain to win as the sampling error in the poll declines. The third candidate on a list with expected share less than $2Q$ is virtually certain to lose (again, as the sampling error declines). And so forth.

Well-informed instrumental voters in this system will avoid voting for a list that has no chance at a seat. They will concentrate instead on lists in competition for the last-allocated remainder seat, where their votes are most likely to improve the outcome. This leads to a prediction of at most $M + 1$ serious lists.

All of these predictions about the equilibrium number of candidates and lists need to be hedged in various ways. Suffice it to say here that they all depend on two key theoretical assumptions – one an assumption about voters' motivations, one an assumption about voters' expectations. Although closely approximated in some empirical instances, enough to make the theory interesting, these assumptions are certainly not universal. Indeed, part of the value of the formal modeling to follow is that it identifies potentially measurable variables that should, in addition to electoral structure, affect the incidence of strategic voting. All this will be elaborated in the appropriate chapters.

2.5 OUTLINE OF PART II

Part II of the book, which commences with the next chapter, considers strategic voting in each of the three main electoral systems that Duverger originally studied, as well as in various systems that he did not study. Chapter 3 sets the stage for the analysis by discussing electoral systems in general, providing a snapshot of the full range of systems in use circa 1992 in democratic elections, and suggesting informally some of the ways in which different features of these systems facilitate strategic voting or (something which is more pertinent to later parts of the book) strategic entry and alliance formation. Chapters 4, 5, and 6 investigate strategic voting in single-member districts operating under several single-ballot voting procedures (4), multimember districts operating under various forms of proportional representation (5), and single-member districts operating under dual-ballot procedures (6). Each of the analytical chapters includes, in addition to comments on the formal model and associated equilibrium results, a review of relevant literature pertaining to the systems under discussion. Two of the chapters (4 and 5) present original empirical evidence as well.

When put together, the picture that emerges from Part II is very much in accord with the Leys-Sartori conjecture. There is a continuum of systems, ranging from those in which strategic voting imposes a constraining upper bound, to those in which it imposes a rarely-constraining or unconstraining upper bound, on the number of parties.

Part II closes with a discussion (in Chapter 7) of two themes that are largely ignored in the institutionalist view of electoral politics. First, I argue that a proper understanding of the institutionalist results implies that the number of candidates/lists that compete in a system will be an interactive product of both social diversity and the permissiveness of the electoral rules. This leads into the empirical analysis of Chapter 11. Second, I note that there are modalities of strategic voting that do not operate to the detriment of small or weak parties, and that some electoral systems promote these kinds of strategic voting rather than the classic "wasted vote" kind.

4

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

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).

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

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 bipartism. 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

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

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 $\pi_i = (\pi_{i1}, \ldots, \pi_{iK})$, 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

[2]This 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).
[3]Voters' 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 $u_i = (u_{i1}, \ldots, u_{iK})$, an element in the set $U = \{(u_1, \ldots, u_K)$: $\max\{u_j\}=1$ & $\min\{u_j\}=0$ & $u_i=u_k$ only if $j = k\}$.
[4]Given 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).

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

[5]Another 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.

(i.e., casting a ballot in $V(u_i;\pi, n)$), 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 $\pi_1 \geq \pi_2 \geq \ldots \geq \pi_K$. 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:

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.

[6]Even 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.

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-

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

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 +/– 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 –

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

[7]Even 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.

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.

[8]There 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.

[9]The 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."

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

Independent variables	*Dependent variable and year of election*			
	FDPLOSS 1987	*FDPLOSS 1990*	*GRLOSS 1987*	*GRLOSS 1990*
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

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

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)

[13]Galbraith 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.

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.

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

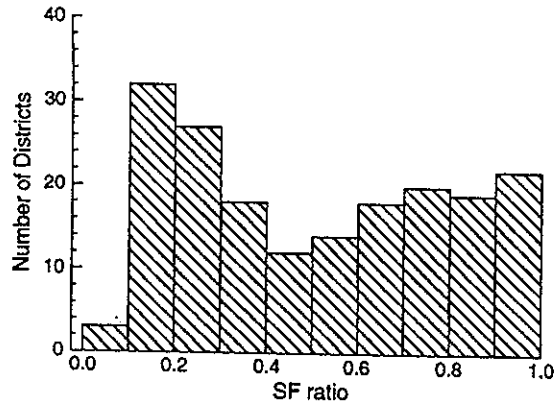

Figure 4.1. Testing the bimodality hypothesis in moderately close districts: British elections, 1983–1992

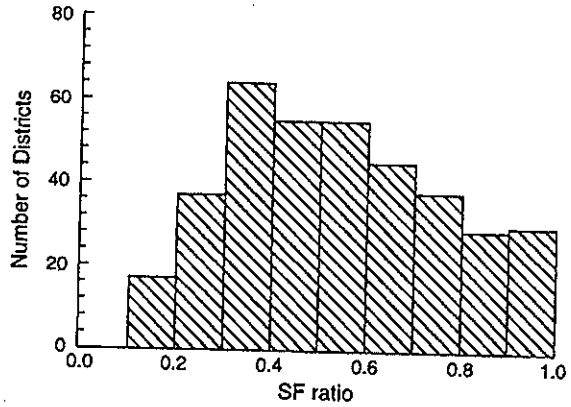

Figure 4.2. Testing the bimodality hypothesis in districts that were not closely contested: British elections, 1983–1992

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-

[14]Note 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.

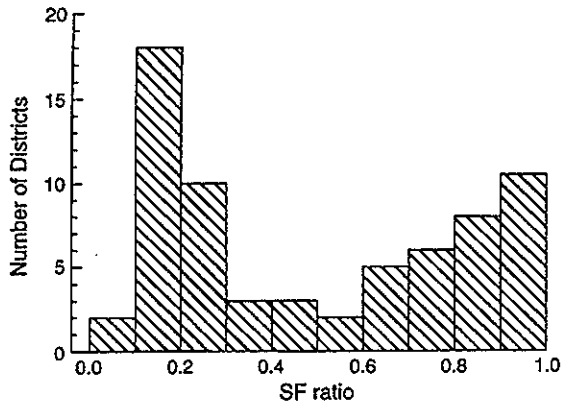

Figure 4.3. Testing the bimodality hypothesis in very close districts: British elections, 1983–1992

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

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

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.

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

[15]Currently, 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).

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

[16]Certain 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).

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),

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

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-

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.

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

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 arguement 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.