# Three Aspects of Multicandidate Competition in Plurality Rule Elections

Thesis by

Mark Fey

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology
Pasadena, California

1995

(Submitted October 17, 1994)

## Dedication

To my family

## Acknowledgments

In completing this project I have relied on the help and support of many people, all of who deserve my thanks. I have also done a lot of hard work, grappled with some challenging topics, and even learned a little bit about myself on the way. My four years in Baxter Hall have truly been a learning experience, in more ways than one.

My advisor, Tom Palfrey, has had the most direct and most valuable effect on the shape and direction of this research. He has always provided enthusiastic and insightful commentary on my ideas and stratagems. If this work has any merit, a great deal of the credit must go to him. The other members of my examination committee also deserve a special thanks. Richard McKelvey was influential in all stages of this work. Mike Alverez encouraged and aided my interest in "real" political science at every turn. Finally, Peter Ordeshook sparked my interest in formal political theory, dating back to PS12. Several other faculty members, though not directly involved in this thesis, were nevertheless significant. Thanks goes to Charlie Plott for helping shape my career path, Scott Page for always being entertaining, Dave Porter for putting me in my place, and Kim Border for answering my endless LATEXquestions.

I have been fortunate to attend graduate school with a great group of fellow students. I am particularly grateful for all the aid and understanding offered by my officemates, John Duggan and Katia Sherstyuk. John was paradoxically both a great distraction and a great inspiration, and helped me sort out the gems in the rough of my many ideas. Additional thanks goes to Charles Noussair, Charles Polk, Dave Schmidt and Pete Coughlan. I am grateful for the helpful staff in Baxter Hall, especially super sports fan Eloisa Imel for her assistance in countless matters, Wes Boudville for installing software and reinstalling Mathematica and Mary Schaffler for

her fluent Klingon. My friends outside of school have also been a big help during this time. Phil, Carver, James and Dwight have all provided relaxation and entertainment.

Earlier drafts of these chapters benefited from comments by Liz Gerber, Morgan Kousser, Tim Feddersen, Roger Myerson, Linda Cohen, Howard Rosenthal, Robin Hanson, and a Caltech proofreader. Financial support is gratefully acknowledged from the Haynes Foundation.

My deepest thanks are reserved for my family and, most of all, my wife Kate. My parents instilled the love of learning in me, without which I would not be where I am today. Kate provides the love of my life, without which I would not want to be who I am. This work represents another milestone in our journey together . . .

#### **Abstract**

This thesis considers three issues relevant to multicandidate competition in plurality rule elections — entry decisions by candidates, strategic voting, and informational concerns.

In the first chapter, we consider a model of spatial competition with entry introduced by Palfrey (1984). In the model, there are two dominant candidates and a potential entrant. The established candidates choose positions first, anticipating the entry decision of the third candidate. In the resulting equilibrium, this threat of entry forces the established parties to adopt spatially separated "moderate" positions. We develop a general model that applies to the complex institutional features of modern elections. Specifically, we introduce the winner-take-all aspects of the Electoral College and show how these characteristics make a difference in the equilibrium predictions of the model. We find that, in one case at least, increasing diversity in the electorate causes the established candidates to initially shift toward more moderate positions and then back toward more extreme positions.

The second chapter examines strategic voting and Duverger's Law. A voter whose favorite candidate has no hope of victory may choose to avoid a "wasted vote" by settling for a less preferred candidate with a higher chance of winning. This behavior erodes the electoral support of minor candidates and results in Duverger's Law: "plurality rule elections favor two party competition." Palfrey (1989) constructs an incomplete information game among voters and shows that as the size of the electorate gets large, the support for the least popular candidate vanishes. We show that there exist equilibria in this model in which all three candidates receive votes under plurality rule, in violation of Duverger's Law, as suggested by Myerson and Weber (1993). However, we proceed to demonstrate that these equilibria are unstable and

any uncertainty by voters leads voters back toward Duvergerian equilibria. In addition, we develop a dynamic model of pre-election polls that describes how voters react to changing information about the viability of the candidates and show that this process leads voters to coordinate on a Duvergerian outcome. Thus, we not only reestablish Duverger's Law, we also describe how voters can use pre-election polls to coordinate on a particular pair of competitive candidates.

In the third chapter we analyze the relationship between voter information and election outcomes in a multicandidate setting. We extend a model originally developed by McKelvey and Ordeshook for two candidate elections to the multicandidate case. In the model, voters are either informed or uninformed about the exact positions of the candidates. The uninformed voters, however, are able to make plausible inferences about these positions based on the vote share each candidate receives. In equilibrium, voters vote optimally, given their beliefs, and beliefs are self-fulfilling in the sense that they are not contradicted by observable information. Our first result is that in the unique voter equilibrium of our model, all voters, informed and uninformed alike, vote as if they had perfect information. We then define a dynamic process involving a sequence of polls that illustrates that this equilibrium is always reached. In addition, we obtain results about candidate positioning equilibria when candidates are also uncertain about the characteristics of the voters. Finally, we show that if a small minority of voters are fully informed and use this information to vote strategically, in equilibrium all voters, including uninformed sincere voters, act as if they were voting strategically based on full information. The uninformed voters view the lack of support for trailing candidates by informed voters as evidence that these candidates are undesirable and react by voting for a more prominent candidate.

## Contents

1	Introduction		
2	Candidate Positions and Institutional Features		7
	2.1	Introduction	7
	2.2	Notation, Assumptions and Equilibrium Concepts	10
	2.3	Equilibrium in a General Setting	22
	2.4	Conclusion	49
	App	endix: Approximation Results	50
3	Sta	bility, Polls, and Duverger's Law	54
	3.1	Introduction	54
	3.2	The Model	59
	3.3	Existence of Non-Duvergerian Equilibria	68
	3.4	Stability and Dynamics of Non-Duvergerian Equilibria	73
	3.5	Polls as a Coordination Device	82
	3.6	Conclusion	85
4	Information, Polls, and Bandwagons		
	4.1	Introduction	87
	4.2	Notation and Definitions	92

## viii

Biblio	graphy	121
4.6	Conclusion	120
4.5	Strategic Voting and Bandwagons	116
4.4	Full Equilibria	108
4.3	Characterization and Convergence of Voter Equilibria	98

## List of Figures

2.1	Established Candidate Positions and Electoral Diversity	46
3.1	$h_n(v)$ for three values of $n$	72
3.2	Sequence of Equilibria Converging to $\tilde{v}$	74
3.3	$h_n(v)$ and the Adjustment Path of Opinion Polls	84
4.1	Inference of Uninformed Voters	97
4.2	Vote Response In Successive Polls	104
4.3	Candidate Beliefs In A Full Equilibrium	111

## List of Tables

1.1	Notable Third-party and Independent Candidates Since 1912	5
3 1	Voter Types in the Electorate	61

## Chapter 1

## Introduction

The study of multicandidate<sup>1</sup> elections, especially via formal models, has for the most part been neglected, especially in the context of American politics. While there are exceptions, both empirical and theoretical research has mainly focused on two candidate elections. This is unfortunate, as multicandidate elections offer a unique set of issues that must be considered. The three issues we address in this work are entry decisions by candidates, strategic voting, and informational concerns. Each topic will be developed fully in the ensuing chapters, but we give a brief overview of each subject here.

The formal literature on multicandidate electoral competition can be divided into three broad categories: models that focus on candidates as strategic actors, models that concentrate on voters as rational actors, and models that include both rational candidates and strategic voters. Models that focus on candidates as strategic actors have a history that traces back to Hotelling's spatial model of dualistic competition in

<sup>&</sup>lt;sup>1</sup>Heeding the advice of Linda Cohen, we note from the outset that, according to Webster's, the prefix "multi" means "more than two." Thus we distinguish between two candidate and multicandidate elections. While on the subject, we also note our use of the closed form "multicandidate" rather than the hyphenated form "multi-candidate." While both usages are commonly seen in the political science literature, we follow *The Chicago Manual of Style*'s recommendation in this matter.

economics and Downs' application of this work to electoral competition. These models describe two candidate competition, and are the basis for the celebrated Median Voter Theorem. Shepsle (1991) reviews the large body of literature, both economic and political, extending models of spatial competition to multiple agents (Eaton and Lipsey 1975; Cox 1987b). These models consider equilibrium strategies of spatial competition of a fixed number of candidates/firms. Palfrey (1984) was the first to consider the effect of entry on political competition. In Palfrey's model, two established candidates choose platforms competitively while anticipating entry of a vote-maximizing third candidate. In the resulting equilibrium, this threat of entry forces the established parties to adopt spatially separated "moderate" positions in reaction. Other authors have relaxed the assumptions of Palfrey's model (Lian 1992; Weber 1992) and extended this approach to multiseat contests (Greenberg and Shepsle 1987).

Whatever the assumptions about the number of candidates or the sequence of actions, these models generally assume the issue space is one-dimensional, usually the unit interval. While this may reasonably describe some electoral institutions, it fails to capture the particular institutional features that are often present in modern electoral systems. American presidential elections, for example, are not national contests, but rather are fifty-one simultaneous state elections in which a candidate must win a plurality in a (weighted) majority of the states in order to win the election outright. If no candidate obtains a majority of electoral votes, Congress is empowered to choose the President. In chapter 2, we extend Palfrey's model of spatial competition with entry to include such institutional features of modern elections. We find that including these characteristics makes a difference in the equilibrium predictions of the model. We find that, in one case at least, increasing diversity in the electorate causes the established candidates to initially shift toward more moderate positions and then back toward more extreme positions.

A second class of formal models of multicandidate elections focus on voters as rational actors. Research has concentrated on issues of the "wasted vote" phenomenon and Duverger's Law. A voter whose favorite candidate has no chance of winning may choose to avoid a "wasted vote" by voting for a less preferred candidate in hopes of perhaps making a difference. If this behavior is widespread in the electorate, minor candidates may find their electoral support eroding away as voters swing their support to more viable candidates. Thus, Duverger's Law — plurality rule ensures two party competition — may be a consequence of strategic calculations by voters in the electorate. We use the term strategic voting<sup>2</sup> to describe this behavior. In political science, strategic voting has received more attention in research on voting on bills and amendments in committee or legislative proceedings which have well defined agenda rules (McKelvey and Niemi 1978; Ordeshook and Palfrey 1988). In the context of mass elections, Cox (1987a) and Palfrey (1989) examine the wasted vote phenomenon in the context of an incomplete information game among voters and show that as the size of the electorate gets large, the support for the least popular candidate vanishes. Thus, they conclude that Duverger was correct and his Law can be considered "proved."

Myerson and Weber (1993) argue that Palfrey has overstated his case in "proving" Duverger's Law. They construct a quasi-Bayesian "voting equilibrium" in which all three candidates receive votes under plurality rule, in violation of Duverger's Law. They also show that this equilibrium is not just a knife-edge case, as claimed by Palfrey. In chapter 3, we attempt to resolve this controversy. We first show that Myerson and Weber's non-Duvergerian equilibria do exist in the model developed by Palfrey. We then argue that these equilibria require extreme coordination and that

<sup>&</sup>lt;sup>2</sup>This is the American terminology. In Britain, "tactical voting" is the normal usage. In addition, "sophisticated voting" is often used in the literature in the context of committees and agendas.

any variation leads voters away from three-party equilibria. Thus, non-Duvergerian equilibria are unstable, while two-party equilibria are not. We develop a dynamic model of pre-election polls that describes how voters react to changing information about the viability of the candidates and show that this process leads voters to coordinate on a Duvergerian outcome. Thus, we not only reestablish Duverger's Law, we also describe how voters can use pre-election polls to coordinate on a particular pair of candidates.

The final topic we address in this work is the information possessed by voters and candidates. This issue is not unique to multicandidate elections, but it is perhaps more relevant. As we note in chapter 4, voters with limited information may have even more difficulty with their decisions when faced with multiple candidates. Chapter 4 is thus concerned with how information is processed and transmitted in multicandidate We extend a model originally developed by McKelvey and Ordeshook for races. two candidate elections to the multicandidate case and obtain several results on the ability of uninformed voters to gather and apply information available to them to their voting decisions. In our model, as in the series of papers by McKelvey and Ordeshook, voters are either informed or uninformed about the exact positions of the candidates. The uninformed voters, however, are able to make plausible inferences about these positions based on the vote share each candidate receives. In equilibrium, voters vote optimally, given their beliefs, and beliefs are self-fulfilling in the sense that they are not contradicted by observable information. We show that in the essentially unique voter equilibrium of our model, all voters, informed and uninformed alike, vote as if they had perfect information. We also define a dynamic process involving a sequence of polls that illustrates that this equilibrium is always reached. In later sections, we introduce candidates who are uncertain about the voters and show that while no three-candidate equilibria exist, a model of third-party entry does have an equilibrium

			Popular	Electoral
Year	Candidate	Party	Vote	Vote
1992	H. Ross Perot	Independent	18.9%	0%
1980	John B. Anderson	${\bf Independent}$	6.6%	0%
1968	George C. Wallace	American Independent	13.5%	8.6%
1948	J. Strom Thurmond	States' Rights Democrat	2.4%	7.3%
1924	Robert M. LaFollete	Progressive	16.6%	2.4%
1912	Theodore Roosevelt	Progressive	27.4%	1.5%

Table 1.1: Notable Third-party and Independent Candidates Since 1912

even with a high degree of uncertainty among the three candidates. Finally, we show that if a small minority of voters are fully informed and use this information to vote strategically, in equilibrium all voters, including uninformed sincere voters, act as if they were voting strategically based on full information. The uninformed voters see the lack of support for trailing candidates by informed voters as evidence that these candidate are undesirable and react by voting for a more prominent candidate. Thus, uninformed voters end up voting strategically even though they are attempting to vote sincerely.

The American multicandidate elections that have received the most attention are presidential elections involving third-party or independent candidates, the most recent example of which is H. Ross Perot in 1992. Table 1.1 lists some notable examples in this century. A better example of multicandidate elections in American politics, though one that has received less attention, occurs in presidential primaries. For as long as direct primaries have been used to select presidential nominees, multicandidate races have been the norm rather than the exception. While there has been some recent excellent explorations of presidential primaries (Bartels 1988; Popkin 1991), it is still an area that invites further investigation.

This research also touches on the widespread use of polls in American politics.

While polls have long been used a tool by political scientists interested in features of the electorate, less attention has been paid to the effect that polls may have in shaping these features. That people are exposed to and pay attention to poll results cannot be denied. According to a recent tabulation of the number of political front-page stories in the Washington Post and the New York Times during the last month of the election in 1988, 42 percent of the Post's and 48 percent of the Times's stories were related to polls. Moreover, polls appeared in 19 percent of the Post's and 9 percent of the Times's headlines (Ratzan 1989). In addition, citizens overwhelmingly report hearing poll results from some source. Gollin (1980) notes that "in two parallel surveys conducted nationally in the heat of the 1976 campaign, only 16 percent failed to recall hearing something in the news or talking with friends' about 'polls showing how candidates for office are doing.' "Thus, this work's finding that voters may able to gather useful information from poll results may have some basis in fact.

Strategic voting is another area that has not received a great deal of attention in political science. Empirical findings have been inconclusive as to the level of strategic voting in the electorate, although evidence that supporters of Perot and Anderson made up their minds much later in the campaign is suggestive. In this area, our findings on the effect of information may have implications on the measurement of strategic voting in the public. In the model presented in chapter 4, based on strategic voting by informed voters, all voters end up voting as if they were strategic. But these decisions are based on incorrect inferences about the quality of the trailing candidates, rather than a conscious decision to avoid a "wasted vote." This could help explain why few voters report voting strategically and yet two candidate competition remains the dominant theme of American politics.

## Chapter 2

## Candidate Positions and

## Institutional Features

### 2.1 Introduction

Models that focus on candidates as strategic actors have a history that traces back to Hotelling's spatial model of dualistic competition in economics and Downs' application of this work to electoral competition. These models describe two candidate competition, and are the basis for the celebrated Median Voter Theorem. This result and its emphasis on "centrality" of candidate positions was taken as a kind of dogma for many years. However, the equilibrium prediction of the Median Voter Theorem, that both candidates adopt policy positions at the median voter's ideal point, invites entry by a third candidate. By picking a policy position near, but not at, the median, an entrant could receive a plurality of votes and thus win the election, assuming that other two candidates' positions remained fixed at the median. Thus, the presence of this third candidate provides an incentive for the two median candidates to change strategies. In fact, it is easy to show that regardless of the configuration of

the three candidates on the policy space, at least one of them will have an incentive to relocate. Put formally, under the usual conditions assumed by the Median Voter Theorem, there is no pure strategy equilibrium for three competing candidates.

In an important paper, Palfrey (1984) establishes the existence of an equilibrium in a model of candidate competition with entry and, in doing so, challenges the centrality results of the Median Voter Theorem. In Palfrey's model, there are two "dominant" or "established" candidates and a prospective entrant. The two dominant candidates choose positions first and then the entrant picks a policy position, knowing the two positions already chosen. The dominant candidates anticipate the possibility of entry, but are not allowed to alter their positions once the third candidate enters the race. In this framework, Palfrey defines a limit equilibrium and proves that such an equilibrium exists. He also shows that this equilibrium involves the two dominant candidates diverging from the position of the median voter in order to head off the entry of a third candidate.

Several papers have extended Palfrey's analysis. Lian (1992) shows that the equilibrium of Palfrey's paper is also the equilibrium of the game in which candidates enter sequentially, instead of having two established candidates move simultaneously, followed by an entrant's move. Weber (1992) generalizes Palfrey's limit equilibrium by defining a hierarchical equilibrium that applies to non-symmetric distributions of voter ideal points. Weber also shows that the strategies of a hierarchical equilibrium correspond to the equilibrium of the sequential move game. Greenberg and Shepsle (1987) examine a model of multiseat contests in which each candidate's objective is to capture one of the seats available in the election. Their model considers an "incumbent" set of candidates faced with the possibility of entry by an additional candidate. An equilibrium in this setting is a set of positions for the incumbent candidates such that an entrant is unable to beat an incumbent candidate for a seat. Unlike Palfrey's

result, Greenberg and Shepsle show that for any district magnitude greater than one, there exists a configuration of voters such that no equilibrium exists. Other related papers include Cox (1987b), which examines equilibria of a wide variety of voting schemes, and Ferejohn and Noll (1988), which allows endogenous entry of candidates, who pay a cost to enter the campaign.

Whatever the assumptions about the number of candidates or the sequence of actions, all of the above models involve a one-dimensional issue space, usually the unit interval, over which a single election is decided. While this may accurately describe some electoral institutions, such as the election of legislators, it does not capture the effect that particular ways of aggregating votes can have. In American presidential elections, for example, a candidate must win a plurality in a (weighted) majority of the states in order to win the election outright. If no candidate obtains a majority of electoral votes, Congress is empowered to choose the President. This chapter extends Palfrey's model of spatial competition with entry to include such institutional features of modern elections.

We begin by reviewing the entry model initially described by Palfrey. In the first section, we define several alternative equilibrium concepts, and show that they are all equivalent in the entry model. Using  $\varepsilon$ -distinguishable equilibrium, we then generalize Palfrey's entry model in order to make further analysis easier. In the next section, we apply this general framework to a model motivated by American presidential elections in which there are several states over which the election is contested. We discover that, in one case at least, increasing diversity in the electorate causes the established candidates to initially shift toward more moderate positions and then back toward more extreme positions. This multiple state model prompts us to consider an alternative view of the voters in the standard model; technical details of this analysis are contained in the appendix.

## 2.2 Notation, Assumptions and Equilibrium Concepts

In this section, we present the mathematical framework of the models we analyze. We begin with the entry model introduced by Palfrey. We then discuss some technical matters regarding the specification of equilibrium. We conclude this section with some additional notation required for the subsequent discussion. As we shall see, the equilibrium predictions of the models we are considering are often critically dependent on the assumptions made. This consideration motivates the analysis in the next section, but until then, we must carefully specify the assumptions of the model.

There are two established or dominant candidates, labeled A and B, and an entrant, labeled C. The two established candidates pick policy positions simultaneously, after which the entrant chooses her<sup>1</sup> policy position. The issue space over which the candidates choose positions is the unit interval [0, 1], and the candidate's positions are denoted  $\theta_A$ ,  $\theta_B$  and  $\theta_C$ , respectively. Voters have symmetric, single-peaked preferences over the issue space and their ideal points are described by a symmetric, unimodal density, with cumulative distribution function F(x).

As mentioned above, we must specify exactly what conditions we place on the distribution of ideal points. Palfrey considers a class of cumulative distributions with six properties. Reorganized and condensed, they are the following:

**Assumption 2.1** F(0) = 0, F(1) = 1, and F(x) is continuous, strictly increasing and twice differentiable on [0,1].

**Assumption 2.2** 
$$F''(x) \ge 0$$
 if  $x < \frac{1}{2}$  and  $F''(x) \le 0$  if  $x > \frac{1}{2}$ .

<sup>&</sup>lt;sup>1</sup>Because we will distinguish between the two established candidates and the entrant, it is convenient to refer to the dominant candidates with male pronouns, and the entrant with a female pronoun.

**Assumption 2.3** 
$$F(x) = 1 - F(1 - x)$$
 for all  $x \in [0, 1]$ .

The first assumption describes a nonatomic distribution with full support. The second assumption requires that F be "essentially" unimodal. In other words, if the distribution has a unique mode, it must occur at  $x = \frac{1}{2}$ . However, distributions with multiple modes are allowed,<sup>2</sup> such as the uniform density on [0,1]. The final assumption requires the distribution function be symmetric, a fairly restrictive requirement, but one that can be relaxed, as we shall see.

Due to the sequential nature of the entrant's decision, the equilibrium in Palfrey's model satisfies a Cournot-Nash conjecture for each dominant candidate and a Stack-elberg conjecture for the entrant. This means that each of the established candidates maximizes his share of the vote, taking the other established candidate's position as fixed, but anticipating the response of the entrant to a different position. In other words, in equilibrium neither candidate A or B will want to shift his position, assuming no reaction from the other established candidate and perfectly forecasting the entrant's response to a shift in his location. Thus, we can describe the equilibrium in terms of the positions of the established candidates, denoting the vote share to candidate k by  $v_k$  and the best response function of the entrant by  $\theta_C$ :

$$\begin{array}{lll} v_A(\theta_A,\theta_B,\theta_C(\theta_A,\theta_B)) & \geq & v_A(\hat{\theta}_A,\theta_B,\theta_C(\hat{\theta}_A,\theta_B)) & \forall \hat{\theta}_A \in [0,1] \\ \\ v_B(\theta_A,\theta_B,\theta_C(\theta_A,\theta_B)) & \geq & v_B(\theta_A,\hat{\theta}_B,\theta_C(\theta_A,\hat{\theta}_B)) & \forall \hat{\theta}_B \in [0,1] \end{array}$$

where the entrant's best response,  $\theta_C(\theta_A, \theta_B)$ , satisfies

$$v_C(\theta_A, \theta_B, \theta_C(\theta_A, \theta_B)) \geq v_C(\theta_A, \theta_B, \hat{\theta}_C) \quad \forall \hat{\theta}_C \in [0, 1]$$

<sup>&</sup>lt;sup>2</sup>See Dharmadhikari and Joag-dev (1988), p.2

Palfrey proves that the unique equilibrium in this model is given by

$$F(\theta_A) = 1 - 2F(1/4 + \theta_A/2)$$
 and  $\theta_B = 1 - \theta_A$ 

In other words, instead of all candidates locating at the median, this model predicts that they will take differentiated positions. They diverge from the median because the established candidates must prevent the entrant from having to much "room" on either side, or in the middle. Thus, the candidates locate in "moderate" positions.

#### Limit Equilibrium

We have ignored several technical details in the above discussion, details that must be handled in our later analysis. Consider the best response function of the entrant,  $\theta_C(\theta_A, \theta_B)$ . First, if there are several distinct positions that will maximize candidate C's vote share, we assume that the entrant picks one at random. Second, there are situations in which  $\theta_C$  is not well-defined. For example, if  $\theta_A = 1/2 = \theta_B$  (the traditional Median Voter Theorem result),  $v_C(1/2, 1/2, 1/2 \pm \alpha) \approx 1/2$  for small  $\alpha$ , but  $v_C(1/2, 1/2, \theta_C = 1/2) = 1/3$ . So the best response of the entrant is to try to enter as close to the median as possible, while not actually entering on the median. In this case,  $\theta_C(1/2, 1/2)$  is not well-defined.

To deal with this type of situation, Palfrey defines a "limit equilibrium," in which the entrant "almost" maximizes votes in a sequence of perturbed games. Formally, if  $v_A(\theta_A, \theta_B, \theta_C)$ ,  $v_B(\theta_A, \theta_B, \theta_C)$ , and  $v_C(\theta_A, \theta_B, \theta_C)$  denote the expected vote to A, B, and C respectively, then for a fixed  $\varepsilon > 0$ , we define the set of  $\varepsilon$ -best responses for the entrant, given positions  $\theta_A$  and  $\theta_B$ , as

$$E_{\varepsilon}(\theta_A, \theta_B) = \{\theta_C \in [0, 1] \mid v_C(\theta_A, \theta_B, \theta_C) > v_C(\theta_A, \theta_B, \hat{\theta}_C) - \varepsilon \quad \forall \hat{\theta}_C \in [0, 1] \}$$

With this definition, Palfrey defines the perturbed game  $\Gamma_{\varepsilon}$  for candidates A and B by specifying the strategy spaces for both candidates as the interval [0,1] and the payoff functions

$$\pi_A^{\varepsilon}(\theta_A, \theta_B) = \int_{E_{\varepsilon}(\theta_A, \theta_B)} v_A(\theta_A, \theta_B, x) \frac{1}{\mu_{\varepsilon}(\theta_A, \theta_B)} dx$$

$$\pi_B^{\varepsilon}(\theta_A, \theta_B) = \int_{E_{\varepsilon}(\theta_A, \theta_B)} v_B(\theta_A, \theta_B, x) \frac{1}{\mu_{\varepsilon}(\theta_A, \theta_B)} dx,$$

where  $\mu_{\varepsilon}(\theta_A, \theta_B) = \int_{E_{\varepsilon}(\theta_A, \theta_B)} dx$  is the (Lebesgue) measure of the entrant's  $\varepsilon$ -best responses. Thus,  $\pi_A^{\varepsilon}(\theta_A, \theta_B)$  and  $\pi_B^{\varepsilon}(\theta_A, \theta_B)$  represent the established candidates' expected vote share given that the entrant is playing a randomly selected  $\varepsilon$ -best response.

Palfrey defines a limit equilibrium as follows:

**Definition 2.1** A limit equilibrium is a pair of policy positions  $(\hat{\theta}_A, \hat{\theta}_B)$  such that

- 1. for every  $\theta_A \neq \hat{\theta}_A$ , there exists a positive  $\bar{\varepsilon}(\theta_A)$  such that  $\varepsilon < \bar{\varepsilon}(\theta_A)$  implies  $\pi_A^{\varepsilon}(\theta_A, \hat{\theta}_B) < \pi_A^{\varepsilon}(\hat{\theta}_A, \hat{\theta}_B)$ , and
- 2. for every  $\theta_B \neq \hat{\theta}_B$ , there exists a positive  $\bar{\varepsilon}(\theta_B)$  such that  $\varepsilon < \bar{\varepsilon}(\theta_B)$  implies  $\pi_B^{\varepsilon}(\hat{\theta}_A, \theta_B) < \pi_B^{\varepsilon}(\hat{\theta}_A, \hat{\theta}_B)$ .

### Hierarchical Equilibrium

Weber (1992) points out that the concept of a limit equilibrium is sensible when the set of best responses for the entrant is empty, but it is less reasonable when a (unique) best response exists. Weber assumes that the entrant chooses this best response when it exists. This slight modification to Palfrey's model also alters the equilibrium concept. Weber calls this modified equilibrium a "hierarchical equilibrium." He shows that a

unique hierarchical equilibrium exists even for non-symmetric distributions of voter ideal points, and that limit equilibria, when they exist, coincide with hierarchical equilibria.

Once again, we must be clear as to the assumptions that Weber invokes. As with Palfrey's model, Weber assumes that the issue space is the closed interval [0, 1], that there is a continuum of voters with symmetric, single-peaked preferences over the issue space, and that voters' ideal points are distributed according to a cumulative distribution function F. Weber assumes that

**Assumption 2.4** F(x) is continuously differentiable, strictly increasing on [0,1] with F(0) = 0 and F(1) = 1.

**Assumption 2.5** There exists (a unique)  $\bar{x} \in [0,1]$  such that F(x) is strictly convex on the interval  $[0,\bar{x}]$  and strictly concave on the interval  $[\bar{x},1]$ .

The first assumption is almost exactly the same as Assumption 2.1 above. The second condition, however, is different from Palfrey's in several respects. Specifically, it allows for non-symmetric single-peaked distributions, but excludes distributions with more than one mode. It thus is more general in some respects and more restrictive in others than Palfrey's assumptions. However, it should be possible to relax the strict single-peakedness requirement and obtain the same results.

In order to define Weber's equilibrium formally, we must first define the set of best responses for candidate C and prove some elementary facts about it. Recall that we denoted the set of  $\varepsilon$ -best responses of the entrant by  $E_{\varepsilon}(\theta_A, \theta_B)$ . We define the set of (strict) best responses for the entrant, given positions  $\theta_A$  and  $\theta_B$  as

$$E_0(\theta_A, \theta_B) = \{ \theta_C \in [0, 1] \mid v_C(\theta_A, \theta_B, \theta_C) \ge v_C(\theta_A, \theta_B, \hat{\theta}_C) \quad \forall \hat{\theta}_C \in [0, 1] \}.$$

Clearly, the following is true:

Lemma 2.1 
$$\bigcap_{\varepsilon>0} E_{\varepsilon}(\theta_A, \theta_B) = E_0(\theta_A, \theta_B)$$

We also collect the following facts about the entrant's payoffs:  $\varepsilon$ -best responses, and (strict) best responses.

**Lemma 2.2** If F(x) satisfies Assumptions 2.4 and 2.5, then for fixed  $\theta_A$  and  $\theta_B$ , the following are true:

- 1.  $v_C(\theta_A, \theta_B, \theta_C)$  is continuous on the intervals  $[0, \theta_A)$ ,  $(\theta_A, \theta_B)$ , and  $(\theta_B, 1]$ .
- 2.  $v_C(\theta_A, \theta_B, \theta_C)$  is increasing on the interval  $[0, \theta_A)$ , and decreasing on the interval  $(\theta_B, 1]$ .
- 3.  $E_0(\theta_A, \theta_B)$  is either empty or a single point in  $[\theta_A, \theta_B]$ .

Proof: To show 1 and 2, we note that for  $\theta_C \in [0, \theta_A)$ ,  $v_C(\theta_A, \theta_B, \theta_C) = F(\frac{\theta_A + \theta_C}{2})$ . As F(x) is continuous and (strictly) increasing, so is  $v_C(\theta_A, \theta_B, \theta_C)$ . A similar argument holds for  $\theta_C \in (\theta_B, 1]$ . Finally, if  $\theta_C \in (\theta_A, \theta_B)$  holds, then  $v_C(\theta_A, \theta_B, \theta_C) = F(\frac{\theta_B + \theta_C}{2}) - F(\frac{\theta_A + \theta_C}{2})$ , which is continuous by the continuity of F(x).

To prove 3, note that as  $v_C$  is strictly increasing on  $[0, \theta_A)$  and strictly decreasing on  $(\theta_B, 1]$ ,  $v_C$  cannot attain a maximum on either of these intervals. Thus,  $\theta_C \in E_0(\theta_A, \theta_B)$  implies  $\theta_C \in [\theta_A, \theta_B]$ . To show there can be only one such point, consider three cases. If  $\frac{\theta_B + \theta_C}{2} > \frac{1}{2}$  and  $\frac{\theta_A + \theta_C}{2} < \frac{1}{2}$ , then  $\theta_C$  satisfies a first-order condition, namely,  $F'(\frac{\theta_B + \theta_C}{2}) - F'(\frac{\theta_A + \theta_C}{2}) = 0$ . The second-order condition is  $\frac{1}{2}[F''(\frac{\theta_B + \theta_C}{2}) - F''(\frac{\theta_A + \theta_C}{2})]$ , which by Assumption 2.5, is strictly negative. So  $\theta_C$  is the unique maximum. If  $\frac{\theta_B + \theta_C}{2}$  and  $\frac{\theta_A + \theta_C}{2}$  are both either bigger or smaller than  $\frac{1}{2}$ , then it is easy to see that  $v_C$  is strictly increasing as  $\theta_C \in [\theta_A, \theta_B]$  moves toward  $\frac{1}{2}$ . Thus, there will either be no unique maximum, or at most one point. We conclude that  $E_0(\theta_A, \theta_B)$  is either empty or a single point in  $[\theta_A, \theta_B]$ .

**Definition 2.2** A pair  $(\theta_A, \theta_B)$  is a hierarchical equilibrium if it constitutes a Nash equilibrium to the following game:

- 1. The strategy set of both players is [0,1].
- 2. The payoff functions of the players are:

$$\nu_A(\theta_A, \theta_B) = \begin{cases} v_A(\theta_A, \theta_B, \theta_C) & \text{if } E_0(\theta_A, \theta_B) = \{\theta_C\} \\ \pi_A(\theta_A, \theta_B) & \text{if } E_0(\theta_A, \theta_B) = \emptyset \end{cases}$$

$$\nu_B(\theta_A, \theta_B) = \begin{cases} v_B(\theta_A, \theta_B, \theta_C) & \text{if } E_0(\theta_A, \theta_B) = \{\theta_C\} \\ \pi_B(\theta_A, \theta_B) & \text{if } E_0(\theta_A, \theta_B) = \emptyset \end{cases}$$

where 
$$\pi_A(\theta_A, \theta_B) = \lim_{\varepsilon \to 0} \pi_A^{\varepsilon}(\theta_A, \theta_B)$$
 and  $\pi_B(\theta_A, \theta_B) = \lim_{\varepsilon \to 0} \pi_B^{\varepsilon}(\theta_A, \theta_B)$ .

In his paper, Weber characterizes the hierarchical equilibrium of the model in a very simple way. This characterization is presented in

**Theorem 2.1 (Weber)** If F(x) satisfies Assumptions 2.4 and 2.5, there is a unique hierarchical equilibrium  $(\theta_A^{\star}, \theta_B^{\star})$ , with  $\theta_A^{\star} < \theta_B^{\star}$ . Moreover, this equilibrium satisfies

$$F(\theta_A^{\star}) = 1 - F(\theta_B^{\star}) = \sup_{\theta_A^{\star} \le \theta_C \le \theta_B^{\star}} v_C(\theta_A^{\star}, \theta_B^{\star}, \theta_C)$$

Proof: See Weber (1992), p. 417.

Weber (1992), Lemma 4.4 shows that  $\lim_{\varepsilon \to 0} \pi_A^{\varepsilon}(\theta_A, \theta_B)$  and  $\lim_{\varepsilon \to 0} \pi_B^{\varepsilon}(\theta_A, \theta_B)$  exist for all pairs  $(\theta_A, \theta_B)$ .

This makes clear the relationship between the entrant's opportunities and the established candidates' strategies. The leftmost term of the characterization is the maximum vote share available to an entrant choosing a position to the left of the dominant candidates. The middle term is the maximum vote share on the right. The rightmost term is the maximum vote share available to a moderate entrant. Thus, a position can only be in equilibrium when the opportunities available to the entrant are balanced: left, middle and right.

#### $\varepsilon$ -distinguishable Equilibrium

An alternative method of making the entrant's best response function well-defined is used by Lian (1992) in his model of sequential entry by candidates. His paper introduces a condition termed " $\varepsilon$ -distinguishability." This condition requires that "[a] candidate who does not move first is required to choose a position at least  $\varepsilon$  away from that of any previous candidate in order to be distinguishable, given  $\varepsilon > 0$  a sufficiently small number."<sup>4</sup>

Since our model has only two stages, we use a modified form of this condition. As the two dominant candidates move simultaneously, it is impossible to prevent them from locating close together or at the same position. But the entrant observes the dominant candidates' positions before she chooses her position, so we can require that " $\varepsilon$ -distinguishability" holds for her choice. We formalize this in

**Assumption 2.6** For a given  $\varepsilon > 0$ ,  $\theta_C$  must satisfy  $|\theta_C - \theta_A| \ge \varepsilon$  and  $|\theta_C - \theta_B| \ge \varepsilon$ .

Clearly this assumption, along with Lemma 2.2, guarantees that the set of entrant best responses,  $\theta_C(\theta_A, \theta_B)$ , is nonempty for all  $\theta_A$  and  $\theta_B$ . However, we must deal with the possibility that  $\theta_C(\theta_A, \theta_B)$  is multivalued. If we assume that Assumption 2.5

<sup>&</sup>lt;sup>4</sup>See Lian (1992), p. 4

holds, then by Lemma 2.2,  $\theta_C(\theta_A, \theta_B)$  can contain at most three distinct points. In fact, we can show

**Lemma 2.3** For any  $\varepsilon > 0$ , if F(x) satisfies Assumptions 2.4 and 2.5 and  $\theta_C(\theta_A, \theta_B)$  satisfies Assumption 2.6, then

$$\theta_C(\theta_A, \theta_B) = L \cup C \cup R$$

where L is either  $\{\theta_A - \varepsilon\}$  or the empty set, R is either  $\{\theta_B + \varepsilon\}$  or the empty set, and C is either some point in  $[\theta_A + \varepsilon, \theta_B - \varepsilon]$  or the empty set, and not all L, C, and R are empty.

Proof: Lemma 2.2, part 1, and Assumption 2.6 insure that we have a continuous function on a compact set, so  $v_C$  attains its maximum somewhere on the set. Thus  $\theta_C(\theta_A, \theta_B)$  is nonempty for all  $\theta_A$  and  $\theta_B$ . By the proof of Lemma 2.2, part 3,  $\theta_C(\theta_A, \theta_B)$  includes at most one point in  $[\theta_A + \varepsilon, \theta_B - \varepsilon]$ . But for  $\theta_C \leq \theta_A - \varepsilon$ ,  $v_C = F((\theta_C + \theta_A)/2)$ , which is strictly increasing by Assumption 2.4. So the only point that could be in  $\theta_C(\theta_A, \theta_B)$  if  $\theta_C \leq \theta_A - \varepsilon$  is  $\{\theta_A - \varepsilon\}$ . Likewise, the only point that could be in  $\theta_C(\theta_A, \theta_B)$  if  $\theta_C \geq \theta_B + \varepsilon$  is  $\{\theta_B + \varepsilon\}$ . So  $\theta_C(\theta_A, \theta_B)$  must be of the form  $L \cup C \cup R$  where L is either  $\{\theta_A - \varepsilon\}$  or the empty set, R is either  $\{\theta_B + \varepsilon\}$  or the empty set, and C is either some point in  $[\theta_A + \varepsilon, \theta_B - \varepsilon]$  or the empty set, and not all L, C, and R are empty.

Actually, this is not as complicated as it looks. In the symmetric case, for example, L and R will be either both empty or both nonempty.

With Lemma 2.3, we can proceed to define an appropriate equilibrium notion for  $\varepsilon$ -distinguishability. As the entrant's set of best responses is nonempty for all  $\varepsilon > 0$ ,

we are able to impose the normal Nash equilibrium requirements for positive  $\varepsilon$ . We thus define an equilibrium to be the limit of the Nash equilibria as  $\varepsilon \to 0$ .

**Definition 2.3** If  $\theta_C(\theta_A, \theta_B)$  satisfies Assumption 2.6, then  $(\theta_A, \theta_B)$  is an equilibrium if  $\theta_A = \lim_{\varepsilon \to 0} \theta_A^{\varepsilon}$  and  $\theta_B = \lim_{\varepsilon \to 0} \theta_B^{\varepsilon}$ , where the pair  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfies, for positive  $\varepsilon$ :

$$\begin{array}{lll} \tilde{v}_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta_C(\theta_A^\varepsilon,\theta_B^\varepsilon) & \geq & \tilde{v}_A(\hat{\theta}_A,\theta_B^\varepsilon,\theta_C(\hat{\theta}_A,\theta_B^\varepsilon)) & \qquad \forall \hat{\theta}_A \in [0,1] \\ \\ \tilde{v}_B(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta_C(\theta_A^\varepsilon,\theta_B^\varepsilon) & \geq & \tilde{v}_B(\theta_A^\varepsilon,\hat{\theta}_B,\theta_C(\theta_A^\varepsilon,\hat{\theta}_B)) & \qquad \forall \hat{\theta}_B \in [0,1] \end{array}$$

where

$$\tilde{v}_k(\theta_A, \theta_B, \theta_C(\theta_A, \theta_B)) = \sum_{\theta \in \theta_C(\theta_A, \theta_B)} \frac{1}{|\theta_C(\theta_A, \theta_B)|} v_k(\theta_A, \theta_B, \theta).$$

#### Equivalence of Equilibrium

Now that we have defined three alternative concepts of equilibrium, we would like to understand the relationship between them. As we show in the next theorem, it turns out that they are equivalent. We show this in

**Theorem 2.2** If F(x) satisfies Assumptions 2.3, 2.4 and 2.5, then the equilibrium of Definition 2.3 coincides with the limit equilibrium specified by Definition 2.1.

Proof: We first prove that a limit equilibrium is a  $\varepsilon$ -distinguishable equilibrium. Let  $(\theta_A, \theta_B)$  be a limit equilibrium. By Palfrey (1984), Theorem 2,  $(\theta_A, \theta_B)$  must be of the form  $F(\theta_A) = 1 - 2F(1/4 + \theta_A/2)$  and  $\theta_B = 1 - \theta_A$ . Now, let  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfy  $F(\theta_A - \varepsilon) = 1 - 2F(1/4 + \theta_A/2)$  and  $\theta_B = 1 - \theta_A$ . First, we must show such a pair exists for sufficiently small  $\varepsilon$ . Let  $g(\theta_A) = 1 - 2F(1/4 + \theta_A/2) - F(\theta_A - \varepsilon)$ ; clearly  $g(\theta)$  is continuous and  $g(\varepsilon) = 1 - 2F(1/4 + \varepsilon/2)$ , which is positive for sufficiently small  $\varepsilon$ , and  $g(1/2) = -F(1/2 - \varepsilon)$ , which is negative for sufficiently small  $\varepsilon$ . Therefore, there

must exist  $\theta_A^{\varepsilon}$  such that  $g(\theta_A^{\varepsilon}) = 0$ . So  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfies  $F(\theta_A - \varepsilon) = 1 - 2F(1/4 + \theta_A/2)$  and  $\theta_B = 1 - \theta_A$ . We must now show that  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfies Definition 2.3. Fix  $\theta_B^{\varepsilon}$ . Notice that the condition  $F(\theta_A - \varepsilon) = 1 - 2F(1/4 + \theta_A/2)$  implies that  $\theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}) = \{\theta_A - \varepsilon\} \cup \{\theta_B + \varepsilon\} \cup \{1/2\}$ . Then

$$\begin{split} \tilde{v}_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta_C(\theta_A^\varepsilon,\theta_B^\varepsilon)) &= \sum_{\theta\in\theta_C(\theta_A^\varepsilon,\theta_B^\varepsilon)} \frac{1}{|\theta_C(\theta_A^\varepsilon,\theta_B^\varepsilon)|} \, v_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta) \\ &= \sum_{\theta\in\{\theta_A^\varepsilon-\varepsilon\}\cup\{\theta_B^\varepsilon+\varepsilon\}\cup\{1/2\}} \frac{1}{3} \, v_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta) \\ &= \frac{1}{3} \, \left[ v_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta_A^\varepsilon-\varepsilon) + v_A(\theta_A^\varepsilon,\theta_B^\varepsilon,\theta_B^\varepsilon+\varepsilon) + v_A(\theta_A^\varepsilon,\theta_B^\varepsilon,1/2) \right] \\ &= \frac{1}{3} \, \left[ (1/2 - F(\theta_A^\varepsilon-\varepsilon/2)) + F((\theta_A^\varepsilon/2 + 1/4)) + 1/2 \right] \\ &= \frac{1}{3} \, \left[ 3/2 - F(\theta_A^\varepsilon-\varepsilon/2) - 1/2F(\theta_A^\varepsilon-\varepsilon) \right] \\ &> \frac{1}{2} \, \left[ 1 - F(\theta_A^\varepsilon-\varepsilon/2) \right]. \end{split}$$

Consider  $\hat{\theta}_A \neq \theta_A^{\varepsilon}$ . There are two cases.

- CASE 1: If  $\hat{\theta}_A < \theta_A^{\varepsilon}$ , then  $F(\hat{\theta}_A \varepsilon) < 1 2F(1/4 + \hat{\theta}_A/2)$ . So the entrant will enter in the center. This means  $\tilde{v}_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) = F((\hat{\theta}_A + \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon}))/2)$ . By the same logic as Palfrey (1984), p. 146, this implies  $\tilde{v}_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) < F(1/4 + \theta_A^{\varepsilon}) = 1/2 1/2F(\theta_A^{\varepsilon} \varepsilon) < \tilde{v}_A(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}, \theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}))$  for sufficiently small  $\varepsilon$ . So candidate A will not shift to the left.
- CASE 2:  $\hat{\theta}_A > \theta_A^{\varepsilon}$ , then  $F(\hat{\theta}_A \varepsilon) > 1 2F(1/4 + \hat{\theta}_A/2)$ . So the entrant will enter to the left of the established candidates. This mean  $\tilde{v}_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) = F((\hat{\theta}_A + \theta_B^{\varepsilon})/2) F(\hat{\theta}_A \varepsilon/2)$ . By the same logic as Palfrey (1984), p. 148, this implies  $\tilde{v}_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) < 1 2F(1/4 + \theta_A^{\varepsilon}) < 1/2 1/2F(\theta_A^{\varepsilon} \varepsilon) < \tilde{v}_A(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}, \theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}))$  for sufficiently small  $\varepsilon$ . So candidate A will not shift to the right.

A similar argument applies to  $\theta_B^{\varepsilon}$ , so  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfies

$$\begin{array}{lll} \tilde{v}_A(\theta_A^{\varepsilon},\theta_B^{\varepsilon},\theta_C(\theta_A^{\varepsilon},\theta_B^{\varepsilon}) & \geq & \tilde{v}_A(\hat{\theta}_A,\theta_B^{\varepsilon},\theta_C(\hat{\theta}_A,\theta_B^{\varepsilon})) & \forall \hat{\theta}_A \in [0,1] \\ \\ \tilde{v}_B(\theta_A^{\varepsilon},\theta_B^{\varepsilon},\theta_C(\theta_A^{\varepsilon},\theta_B^{\varepsilon}) & \geq & \tilde{v}_B(\theta_A^{\varepsilon},\hat{\theta}_B,\theta_C(\theta_A^{\varepsilon},\hat{\theta}_B)) & \forall \hat{\theta}_B \in [0,1] \end{array}$$

where  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfies  $F(\theta_A - \varepsilon) = 1 - 2F(1/4 + \theta_A/2)$  and  $\theta_B = 1 - \theta_A$ . The defining condition implies  $\lim_{\varepsilon \to 0} \theta_A^{\varepsilon}$  and  $\lim_{\varepsilon \to 0} \theta_B^{\varepsilon}$  are well-defined and, moreover,  $\lim_{\varepsilon \to 0} \theta_A^{\varepsilon} = \theta_A$  and  $\lim_{\varepsilon \to 0} \theta_B^{\varepsilon} = \theta_B$ . So, we can conclude that  $(\theta_A, \theta_B)$  is a  $\varepsilon$ -distinguishable equilibrium.

Next, we show that a  $\varepsilon$ -distinguishable equilibrium is a limit equilibrium. Suppose  $(\theta_A, \theta_B)$  is a  $\varepsilon$ -distinguishable equilibrium and assume that  $(\theta_A, \theta_B)$  is not a limit equilibrium. By Assumption 2.3, it must be the case that  $\theta_B = 1 - \theta_A$ . Therefore, we concentrate on candidate A and omit the analogous argument for candidate B. If  $(\theta_A, \theta_B)$  is not a limit equilibrium, then there exists a  $\hat{\theta}_A \neq \theta_A$  such that for all positive  $\varepsilon$ ,  $\pi_A^{\varepsilon}(\hat{\theta}_A, \theta_B) > \pi_A^{\varepsilon}(\theta_A, \theta_B)$ . As  $(\theta_A, \theta_B)$  is a  $\varepsilon$ -distinguishable equilibrium, then for all positive  $\varepsilon$ ,  $\tilde{v}_A(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}, \theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}) \geq \tilde{v}_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon}))$ . But by Assumption 2.3 and Weber (1992), Lemma 4.4,  $\lim_{\varepsilon \to 0} \pi_A^{\varepsilon} = \lim_{\varepsilon \to 0} \tilde{v}_A$ . But this implies  $\lim_{\varepsilon \to 0} \tilde{v}_A(\hat{\theta}_A, \theta_B) = \lim_{\varepsilon \to 0} \pi_A^{\varepsilon}(\hat{\theta}_A, \theta_B) > \lim_{\varepsilon \to 0} \pi_A^{\varepsilon}(\theta_A, \theta_B) = \lim_{\varepsilon \to 0} \tilde{v}_A(\hat{\theta}_A, \theta_B)$ . But this means  $(\theta_A, \theta_B)$  cannot be a  $\varepsilon$ -distinguishable equilibrium. This is a contradiction, thus we conclude that  $(\theta_A, \theta_B)$  is a limit equilibrium.

This theorem, coupled with Theorem 2.8 of Weber's paper, proves the equivalence of the three equilibrium concepts discussed in this chapter: limit equilibrium, hierarchical equilibrium, and  $\varepsilon$ -distinguishable equilibrium. This is important because the last type mentioned is much easier to evaluate than the other two. We now know that it is legitimate to use  $\varepsilon$ - distinguishable equilibrium in our further analysis, which we now do.

## 2.3 Equilibrium in a General Setting

Even if we restrict our attention to simple one-dimensional issue spaces, we find that our equilibrium concepts and outcomes depend crucially on the specification of voter ideal points. Needless to say, when we attempt to add institutional features that would give our model a more "realistic" feel, the difficulties only multiply. Equilibrium outcomes then depend on several more layers of assumptions and specifications. Up until now, each model must be examined and appropriate equilibria solved for individually.

Next suppose there are three groups of voters, called "states," over which the election is held. As is the case in American presidential elections, we suppose that a particular candidate "wins a state" if the candidate receives a plurality of the vote in that state, and the winner of the election is the candidate who wins a majority of the states. In other words, we assume that each state has one state-vote, which is awarded to the plurality winner of that state; the winner of the election is the candidate who receives a majority of the state-votes.<sup>5</sup> In keeping with our connection to American presidential elections, these "state elections" all occur simultaneously and candidates' issue positions cannot vary across states.

In this case, the usual technique would involve specifying the candidate objective functions on one or all of the states, the distribution of voter ideal points on the states, the relation between candidate positions and votes received in a particular state, and the method of aggregation of the vote shares in the states to determine the outcome of the election.

We simplify the analysis in the following way. First, we assume that candidates

<sup>&</sup>lt;sup>5</sup>In case no candidate wins a majority, a tie-breaking procedure, which will not be specified here, comes into play.

maximize the probability that they win the election. We exclude from consideration each candidate's preferences over the size of his electoral support or situations in which the candidate is not elected, for example. Note that this is equivalent to requiring candidates to maximize their utility, where we assume the utility of winning the election is greater than the utility of losing the election. Second, we will ignore the specific institutional features of the model and instead look directly at the probability that a candidate k wins the election. We denote this probability, for a particular choice of candidate positions, by  $p_k(\theta_A, \theta_B, \theta_C)$ . We will give conditions on this function that guarantee that an equilibrium of a certain form exists.

Before we specify these conditions, we must establish the framework of the model. We will either generalize, or use directly, the features of the entry model discussed in the previous section. Specifically, we assume that

- There are two dominant candidates, labeled A and B, and an entrant, labeled C.
- The established candidates pick policy positions simultaneously, after which the entrant chooses her policy position.
- The policy space is  $\Theta$ , a closed interval  $[\underline{\theta}, \overline{\theta}]$  with midpoint  $\frac{1}{2}$ .
- The candidates' positions are denoted  $\theta_A$ ,  $\theta_B$  and  $\theta_C$ , where these are all points in  $\Theta$  and, for clarity,  $\theta_A \leq \theta_B$ .
- Voters have symmetric, single-peaked preferences over the issue space I, which need not be equal to  $\Theta$ .
- A candidate's probability of winning the election is a function of the candidates' positions only. Formally, there exists a real-valued function for each candidate

$$k,\,p_k(\theta_A,\theta_B,\theta_C)$$
 :  $\Theta\times\Theta\times\Theta\to[0,1]$  such that  $\sum p_k\leq 1.^6$ 

The last assumption listed is the sharpest departure from Palfrey's model of entry. We will focus on it as we list conditions for the existence of equilibrium.

#### **Equilibrium Conditions**

In order to characterize equilibria in this general framework, we begin by analyzing the entrant's decision problem.

**Assumption 2.7 (Continuity)** For any  $\theta_A$ ,  $\theta_B \in \Theta$ ,  $p_C(\theta_A, \theta_B, \cdot)$  is continuous on the intervals  $[\underline{\theta}, \theta_A), (\theta_A, \theta_B)$ , and  $(\theta_B, \overline{\theta}]$ .

Clearly, this is similar to Lemma 2.2. This is a fairly strong assumption, but we shall see that it is satisfied in a number of useful cases.

In order to simplify the analysis, we must introduce some additional notation for the entrant's maximization. Define, for any  $\varepsilon > 0$ ,

$$C_{\varepsilon}(\theta_A, \theta_B) = \max_{\{\theta_C: \theta_A + \varepsilon < \theta_C < \theta_B - \varepsilon\}} p_C(\theta_A, \theta_B, \theta_C).$$

That is,  $C_{\varepsilon}(\theta_A, \theta_B)$  is the largest probability of winning the election that the entrant can assure herself by entering the campaign between the two established candidates and satisfying the  $\varepsilon$ -distinguishability requirement. Similarly, for any  $\varepsilon > 0$ , define

$$L_{\varepsilon}(\theta_A, \theta_B) = \max_{\{\theta_C: \theta_C \le \theta_A - \varepsilon\}} p_C(\theta_A, \theta_B, \theta_C)$$

$$R_{\varepsilon}(\theta_A, \theta_B) = \max_{\{\theta_C: \theta_C \ge \theta_B + \varepsilon\}} p_C(\theta_A, \theta_B, \theta_C).$$

<sup>&</sup>lt;sup>6</sup>We allow  $\sum p_k < 1$  to handle ties, as discussed below.

Thus,  $L_{\varepsilon}(\theta_A, \theta_B)$  and  $R_{\varepsilon}(\theta_A, \theta_B)$  represent the largest probabilities possible if the entrant enters on the left or right of the established candidates.

By Assumption 2.7,  $C_{\varepsilon}$ ,  $L_{\varepsilon}$ , and  $R_{\varepsilon}$ , are well-defined. Moreover, the maximum of a continuous function is continuous, so  $\lim_{\varepsilon \to 0} C_{\varepsilon}$  exists. So let

$$\lim_{\varepsilon \to 0} C_{\varepsilon}(\theta_A, \theta_B) = C(\theta_A, \theta_B)$$

$$\lim_{\varepsilon \to 0} L_{\varepsilon}(\theta_A, \theta_B) = L(\theta_A, \theta_B)$$

$$\lim_{\varepsilon \to 0} R_{\varepsilon}(\theta_A, \theta_B) = R(\theta_A, \theta_B).$$

It will also be useful to identify the entrant's probability maximizing position. So, define

$$c_{\varepsilon}(\theta_A, \theta_B) = \underset{\{\theta_C: \theta_A + \varepsilon \leq \theta_C \leq \theta_B - \varepsilon\}}{\operatorname{argmax}} p_C(\theta_A, \theta_B, \theta_C).$$

This gives the entrant's best response function,  $\theta_C(\theta_A, \theta_B)$ , given that she enters the campaign in the "middle." Similar expressions are defined for entry on either wing.

$$l_{\varepsilon}(\theta_{A}, \theta_{B}) = \underset{\{\theta_{C}: \theta_{C} \leq \theta_{A} - \varepsilon\}}{\operatorname{argmax}} p_{C}(\theta_{A}, \theta_{B}, \theta_{C})$$
$$r_{\varepsilon}(\theta_{A}, \theta_{B}) = \underset{\{\theta_{C}: \theta_{C} \geq \theta_{B} + \varepsilon\}}{\operatorname{argmax}} p_{C}(\theta_{A}, \theta_{B}, \theta_{C}).$$

Following Lemma 2.3, we assume that  $c_{\varepsilon}$ ,  $l_{\varepsilon}$ , and  $r_{\varepsilon}$  are singletons.

Next we include some symmetry assumptions to capture the symmetric nature of the voting models.

Assumption 2.8 (Symmetry) For any  $\theta_A < \theta_B$ , we assume

1. 
$$C_{\varepsilon}(\theta_A, \theta_B) = C_{\varepsilon}(1 - \theta_B, 1 - \theta_A)$$
.

2. 
$$L_{\varepsilon}(\theta_A, \theta_B) = R_{\varepsilon}(1 - \theta_B, 1 - \theta_A)$$
.

3. 
$$c_{\varepsilon}(\theta_A, \theta_B) = 1 - c_{\varepsilon}(1 - \theta_B, 1 - \theta_A)$$
.

4. 
$$l_{\varepsilon}(\theta_A, \theta_B) = 1 - r_{\varepsilon}(1 - \theta_B, 1 - \theta_A)$$
.

As above, it is immediate that

$$C(\theta_A, \theta_B) = C(1 - \theta_B, 1 - \theta_A)$$

$$L(\theta_A, \theta_B) = R(1 - \theta_B, 1 - \theta_A).$$

We are now ready to define the conditions that are sufficient for the existence of an equilibrium.

#### Condition 2.1 $p_C(\theta_A, \theta_B, \theta_C)$ satisfies

- 1. For any  $\bar{\theta}_B \in \Theta$ ,  $C(\cdot, \bar{\theta}_B)$  is continuous and strictly decreasing on  $[0, \bar{\theta}_B]$ .
- 2. For any  $\bar{\theta}_A \in \Theta$ ,  $C(\bar{\theta}_A, \cdot)$  is continuous and strictly increasing on  $[\bar{\theta}_A, 1]$ .
- 3. If  $\theta_A = \theta_B$ , then  $C(\theta_A, \theta_B) = 0$ .
- 4.  $C(\underline{\theta}, \overline{\theta}) > 0$ .

This condition requires that the entrant's probability of winning when she enters in the middle must shrink as the established candidates move closer together. It also requires that the probability of winning be positive when the established candidates are extremely far apart, and zero when the dominant candidates adopt the same position.

### Condition 2.2 $p_C(\theta_A, \theta_B, \theta_C)$ satisfies

1. For any  $\bar{\theta}_B \in \Theta$ ,  $L(\cdot, \bar{\theta}_B)$  is independent of  $\bar{\theta}_B$ , continuous, and strictly increasing on  $[0, \bar{\theta}_B]$ .

2. For any  $\bar{\theta}_A \in \Theta$ ,  $R(\bar{\theta}_A, \cdot)$  is independent of  $\bar{\theta}_A$ , continuous, and strictly decreasing on  $[\bar{\theta}_A, 1]$ .

3. If 
$$\theta_A = \theta_B > 0$$
, then  $L(\theta_A, \theta_B) > 0$ .

- 4.  $L(\underline{\theta}, \overline{\theta}) = 0$ .
- 5. If  $\theta_A = \theta_B < 0$ , then  $R(\theta_A, \theta_B) > 0$ .
- 6.  $R(\theta, \overline{\theta}) = 0$ .

This condition requires that if the entrant enters to the left, say, of the established candidates, then her probability of winning the election must increase as the dominant candidates move to the right. It also requires there to be a zero probability of winning from an extremist position when the dominant candidates are already extremists and a positive probability of winning with an extremist position when the two major candidates adopt the same moderate position.

**Lemma 2.4** If  $p_C(\theta_A, \theta_B, \theta_C)$  satisfies Assumptions 2.7 and 2.8 and Conditions 2.1 and 2.2, then there exists a unique pair  $(\theta_A^*, \theta_B^*)$  satisfying

$$L(\theta_A^*, \theta_B^*) = C(\theta_A^*, \theta_B^*) = R(\theta_A^*, \theta_B^*).$$

Proof: First, notice that if such a pair  $(\theta_A^*, \theta_B^*)$  exists, then by Assumption 2.8,  $L(\theta_A^*, \theta_B^*) = R(1-\theta_B^*, 1-\theta_A^*) = R(\theta_A^*, \theta_B^*)$ . This implies that  $\theta_A^* + \theta_B^* = 1$ , which means that there exists  $\theta^* \in [0, \overline{\theta} - 1/2]$  such that  $\theta_A^* = 1/2 - \theta^*$  and  $\theta_B^* = 1/2 + \theta^*$ . So consider  $\check{\theta}_A = 1/2 - \check{\theta}$  and  $\check{\theta}_B = 1/2 + \check{\theta}$ . Define  $g(\check{\theta}) = C(1/2 - \check{\theta}, 1/2 + \check{\theta}) - L(1/2 - \check{\theta}, 1/2 + \check{\theta})$ . Notice that if  $g(\check{\theta}) = 0$ , then  $(\check{\theta}_A, \check{\theta}_B)$  satisfies  $L(\check{\theta}_A, \check{\theta}_B) = C(\check{\theta}_A, \check{\theta}_B)$ , which implies by Assumption 2.8 that  $C(\check{\theta}_A, \check{\theta}_B) = R(\check{\theta}_A, \check{\theta}_B)$ . So  $g(\check{\theta}) = 0$  is required for  $(\theta_A^*, \theta_B^*)$ . By Condition 2.1, part 1 and 2, and Condition 2.2, part 1,  $g(\check{\theta})$  is continuous and

strictly increasing on  $[0, \overline{\theta} - 1/2]$ . By Condition 2.1, part 3, and Condition 2.2, part 3, g(0) is negative and by Condition 2.1, part 4, and Condition 2.2, part 4,  $g(\overline{\theta} - 1/2)$  is positive. Thus, by the Intermediate Value Theorem, there exists a unique  $\check{\theta}$  such that  $g(\check{\theta}) = 0$ . Thus, we conclude that there exists a unique pair  $(\theta_A^*, \theta_B^*)$  satisfying  $L(\theta_A^*, \theta_B^*) = C(\theta_A^*, \theta_B^*) = R(\theta_A^*, \theta_B^*)$ .

Note that we did not explicitly use part 2, 4, and 5 from Condition 2.2 in the proof of Lemma 2.4. In fact, these parts are implied by part 1, 3, and 4 of Condition 2.2 and Assumption 2.8.

We now describe the conditions for the established candidates under which the pair identified in Lemma 2.4 is in fact an equilibrium.

# Condition 2.3 When $\theta_B^*$ satisfies Lemma 2.4, $p_A(\theta_A, \theta_B, \theta_C)$ satisfies

1.  $p_A(\cdot, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B))$  is strictly increasing on  $[0, \bar{\theta}_B]$ , and  $p_A(\cdot, \bar{\theta}_B, l_{\varepsilon}(\theta_A, \bar{\theta}_B))$  is strictly decreasing on  $[0, \bar{\theta}_B]$ , for all  $\bar{\theta}_B$  in some neighborhood of  $\theta_B^*$ .

2. 
$$p_A(\theta_A^*, \theta_B^*, c_{\varepsilon}(\theta_A^*, \theta_B^*) \le p_A(\theta_A^*, \theta_B^*, l_{\varepsilon}(\theta_A^*, \theta_B^*))$$
.

Part 1 means that if the entrant is going to choose a policy between the two established candidates, then candidate A's probability of winning rises as he shifts further toward the right; if the entrant is going to choose a policy to the left of the two established candidates, then candidate A's probability of winning rises as he shifts further toward the left. Part 2 implies that if candidate A is located at what we will show is the equilibrium position, then he is at least as well off when the entrant enters in the middle as when she enters on the left.

When Condition 2.3 holds, by symmetry, the following condition should hold.

## Condition 2.4 When $\theta_A^*$ satisfies Lemma 2.4, $p_B(\theta_A, \theta_B, \theta_C)$ satisfies

- 1. For all  $\bar{\theta}_A$  in some neighborhood of  $\theta_A^*$ ,  $p_B(\bar{\theta}_A, \cdot, c_{\varepsilon}(\bar{\theta}_A, \theta_B))$  is strictly decreasing on  $[\bar{\theta}_A, 1]$ , and  $p_B(\bar{\theta}_A, \cdot, r_{\varepsilon}(\bar{\theta}_A, \theta_B))$  is strictly increasing on  $[\bar{\theta}_A, 1]$ .
- 2.  $p_B(\theta_A^*, \theta_B^*, c_{\varepsilon}(\theta_A^*, \theta_B^*) \le p_B(\theta_A^*, \theta_B^*, r_{\varepsilon}(\theta_A^*, \theta_B^*)).$

The following theorem proves that these conditions guarantee the existence of a  $\varepsilon$ -distinguishable equilibrium.

**Theorem 2.3** If  $p_C(\theta_A, \theta_B, \theta_C)$  satisfies Assumptions 2.7 and 2.8 and Conditions 2.1 and 2.2,  $p_A(\theta_A, \theta_B, \theta_C)$  satisfies Condition 2.3, and  $p_B(\theta_A, \theta_B, \theta_C)$  satisfies Condition 2.4, then the unique pair  $(\theta_A^*, \theta_B^*)$  identified in Lemma 2.4 is a  $\varepsilon$ -distinguishable equilibrium as defined in Definition 2.3.

Proof: To see that this is an equilibrium, let  $(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$  satisfy  $L(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}) = C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}) = R(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$ . That such a pair exists follows by an argument precisely like the proof of Lemma 2.4. In fact, it is true that  $\theta_A^{\varepsilon} = 1/2 - \check{\theta}^{\varepsilon}$  and  $\theta_B^{\varepsilon} = 1/2 + \check{\theta}^{\varepsilon}$ . Moreover,  $\lim_{\varepsilon \to 0} \theta_A^{\varepsilon} = \theta_A^*$  and  $\lim_{\varepsilon \to 0} \theta_B^{\varepsilon} = \theta_B^*$  Now consider candidate A. We must show that as  $\varepsilon$  goes to 0, he will have no incentive to deviate from  $\theta_A^{\varepsilon}$ . Fix  $\varepsilon > 0$  and consider  $\hat{\theta_A} \neq \theta_A^{\varepsilon}$ . There are two cases.

- CASE 1:  $\hat{\theta_A} < \theta_A^{\varepsilon}$ . This implies, by the first parts of Conditions 2.1 and 2.2, that  $L(\hat{\theta_A}, \theta_B^{\varepsilon}) < C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$ . Therefore the entrant will enter in the center. But Condition 2.3, part 1, shows that  $p_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) < p_A(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}, \theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}))$  for sufficiently small  $\varepsilon$ . Thus, candidate A will not want to shift to the left.
- CASE 2:  $\hat{\theta_A} > \theta_A^{\varepsilon}$ . This implies, by the first parts of Conditions 2.1 and 2.2, that  $L(\hat{\theta_A}, \theta_B^{\varepsilon}) > C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon})$ . Therefore the entrant will enter on the left. But Condition 2.3, part 1, shows that  $p_A(\hat{\theta}_A, \theta_B^{\varepsilon}, \theta_C(\hat{\theta}_A, \theta_B^{\varepsilon})) < p_A(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}, \theta_C(\theta_A^{\varepsilon}, \theta_B^{\varepsilon}))$  for sufficiently small  $\varepsilon$ . Thus, candidate A will not want to shift to the right.

Therefore,  $(\theta_A^*, \theta_B^*)$  is a  $\varepsilon$ -distinguishable equilibrium.

#### Application to the Entry Model

In order to illustrate the application of the above analysis, we return to the entry model described by Palfrey and formalized in Assumptions 2.1, 2.2, and 2.3. By Theorem 2.3, in order to rediscover the equilibrium, we must only check that Assumptions 2.7 and 2.8 and Conditions 2.1, 2.2, 2.3, and 2.4 hold for  $p_C(\theta_A, \theta_B, \theta_C)$ ,  $p_A(\theta_A, \theta_B, \theta_C)$ , and  $p_B(\theta_A, \theta_B, \theta_C)$  respectively.

Before we proceed, however, we must determine what the probability of winning the election is for each candidate. Surprisingly, however, this is not immediately clear. Recall that Palfrey's entry model assumes there is a continuum of voters with ideal points distributed according to F(x). But if we insist on plurality rule, then by Palfrey (1984), Corollary 2 (which says that the entrant never receives a plurality in equilibrium)  $p_C = 0$  in equilibrium. This makes several of our conditions vacuous. It also raises questions as to why the entrant would enter the race, knowing that her probability of winning would be zero. A partially satisfactory reply is that the entrant must only threaten to enter the race in order to force the established candidates to their equilibrium positions. This answer begs the question of the incredible threat that the entrant is making, as well as preferences the entrant has over the established candidates' positions, something clearly not included in the model.

An alternative but mathematically completely identical interpretation is that instead of an infinite number of voters, there is only a single voter with an ideal point that is not known by the candidates. All the candidates know is that the point is distributed according to F(x). In this view, the "vote share" of a candidate is precisely the probability the candidate wins the election. Thus, vote share maximization in Palfrey's model is equivalent to maximizing probability of winning the election. On empirical grounds, this has its flaws. In real elections, third party candidates almost

never win; this does not agree with the equilibrium prediction of the model.

We shall return to this critique later, but for now, we will adopt the second interpretation, and apply the conditions and theorem of the previous section. With the assumptions of the this section, as well as Assumption 2.6, the probabilities of winning are as follows:

$$p_A(\theta_A, \theta_B, \theta_C) = v_A(\theta_A, \theta_B, \theta_C) = \begin{cases} F(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ F(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_A < \theta_B < \theta_C \\ F(\frac{\theta_A + \theta_B}{2}) - F(\frac{\theta_C + \theta_A}{2}) & \text{if } \theta_C < \theta_A < \theta_B \\ \frac{1}{2}F(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A = \theta_B < \theta_C \\ \frac{1}{2}(1 - F(\frac{\theta_C + \theta_A}{2})) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$p_B(\theta_A, \theta_B, \theta_C) = v_B(\theta_A, \theta_B, \theta_C) = \begin{cases} 1 - F(\frac{\theta_C + \theta_B}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_C < \theta_A < \theta_B \\ F(\frac{\theta_B + \theta_C}{2}) - F(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_A < \theta_B < \theta_C \\ \frac{1}{2}F(\frac{\theta_B + \theta_C}{2}) & \text{if } \theta_A = \theta_B < \theta_C \\ \frac{1}{2}(1 - F(\frac{\theta_C + \theta_B}{2})) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$p_C(\theta_A, \theta_B, \theta_C) = v_C(\theta_A, \theta_B, \theta_C) = \begin{cases} F(\frac{\theta_C + \theta_A}{2}) & \text{if } \theta_C < \theta_A \le \theta_B \\ F(\frac{\theta_C + \theta_B}{2}) - F(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F(\frac{\theta_B + \theta_C}{2}) & \text{if } \theta_A \le \theta_B < \theta_C. \end{cases}$$

Now we proceed to check that the appropriate conditions and assumptions are satisfied. In particular, Assumption 2.7 follows from Lemma 2.2 and Assumption 2.8 immediately follows from Assumption 2.3. While parts 3 and 4 of Condition 2.1 are

obviously true, it takes a little more work to show that parts 1 and 2 hold.

#### Lemma 2.5 Condition 2.1 holds.

*Proof*: To show that part 1 holds, recall that

$$C(\theta_A, \theta_B) = \lim_{\varepsilon \to 0} C_{\varepsilon}(\theta_A, \theta_B).$$

We will first investigate the extreme case where the established candidates are on the same side of the midpoint. For any  $\varepsilon > 0$ , since F(x) is symmetric about  $\frac{1}{2}$ , the entrant will locate as close to  $\frac{1}{2}$  as possible when both established candidates are on the same side of the midpoint. Formally,

$$c_{\varepsilon}(\theta_A, \theta_B) = \begin{cases} \theta_B - \varepsilon & \text{if } \theta_A < \theta_B < \frac{1}{2} \\ \theta_A + \varepsilon & \text{if } \frac{1}{2} < \theta_A < \theta_B. \end{cases}$$

By definition (and the fact that  $c_{\varepsilon}$  is single-valued), it must be the case that  $C_{\varepsilon}(\theta_A, \theta_B) = p_C(\theta_A, \theta_B, c_{\varepsilon}(\theta_A, \theta_B))$ . Thus,

$$C_{\varepsilon}(\theta_A, \theta_B) = \begin{cases} F(\theta_B - \frac{\varepsilon}{2}) - F(\frac{\theta_A + \theta_C - \varepsilon}{2}) & \text{if } \theta_A < \theta_B < \frac{1}{2} \\ F(\frac{\theta_A + \theta_B + \varepsilon}{2}) - F(\theta_A + \frac{\varepsilon}{2}) & \text{if } \frac{1}{2} < \theta_A < \theta_B. \end{cases}$$

As F(x) is continuous, when we take limits as  $\varepsilon$  goes to zero, we get:

$$C(\theta_A, \theta_B) = \begin{cases} F(\theta_B) - F(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A < \theta_B < \frac{1}{2} \\ F(\frac{\theta_A + \theta_B}{2}) - F(\theta_A) & \text{if } \frac{1}{2} < \theta_A < \theta_B. \end{cases}$$

So suppose  $\frac{1}{2} < \theta_A < \theta_B$ . Then  $\partial C(\theta_A, \theta_B)/\partial \theta_A = 1/2F'(\frac{\theta_A + \theta_B}{2}) - F'(\theta_A)$ . As F(x) is strictly concave on  $[\frac{1}{2}, 1]$ ,  $F'(\frac{\theta_A + \theta_B}{2}) < F'(\theta_A)$ . Thus,  $\partial C(\theta_A, \theta_B)/\partial \theta_A < 0$ . A similar

argument applies to the case  $\theta_A < \theta_B < \frac{1}{2}$ . So we are left with analyzing the situation when  $\theta_A < \frac{1}{2} < \theta_B$ . In this case, when  $\theta_B$  is fixed,

$$C(\theta_A, \theta_B) = \max_{\theta_A < \theta_C < \theta_B} F(\frac{\theta_C + \theta_B}{2}) - F(\frac{\theta_A + \theta_C}{2}).$$

The maximum must satisfy the first-order condition. In other words, if  $\theta_C^*$  maximizes the function, then

$$F'(\frac{\theta_C^* + \theta_B}{2}) - F'(\frac{\theta_A + \theta_C^*}{2}) = 0.$$

So we can write  $C(\theta_A, \theta_B) = F(\frac{\theta_C^* + \theta_B}{2}) - F(\frac{\theta_A + \theta_C^*}{2})$ . Taking a partial derivative gives

$$\partial C(\theta_A, \theta_B)/\partial \theta_A = F'(\frac{\theta_C^* + \theta_B}{2})(1/2)(\partial \theta_C^*/\partial \theta_A) - F'(\frac{\theta_A + \theta_C^*}{2})(1/2)(1 + \partial \theta_C^*/\partial \theta_A)$$

$$= (1/2)(\partial \theta_C^*/\partial \theta_A)[F'(\frac{\theta_C^* + \theta_B}{2}) - F'(\frac{\theta_A + \theta_C^*}{2})] - (1/2)F'(\frac{\theta_A + \theta_C^*}{2})$$

$$= 0 - (1/2)F'(\frac{\theta_A + \theta_C^*}{2})$$

$$< 0.$$

Thus, in all three cases,  $\partial C(\theta_A, \theta_B)/\partial \theta_A < 0$ . So  $C(\cdot, \bar{\theta}_B)$  is strictly decreasing. A similar argument applied to candidate B shows that  $C(\bar{\theta}_A, \cdot)$  is strictly increasing. So Condition 2.1 holds.

It is much easier to show Condition 2.2 holds. If the entrant chooses a position to the left of the established candidate, then

$$p_C(\theta_A, \theta_B, \theta_C) = F(\frac{\theta_C + \theta_A}{2}).$$

Since this is increasing in  $\theta_C$ ,  $l_{\varepsilon}$  must equal  $\theta_A - \varepsilon$ . So  $l(\theta_A, \theta_B) = \theta_A$  and

$$L(\theta_A, \theta_B) = F(\theta_A).$$

It is immediate that Condition 2.2 holds.

Finally, to confirm Conditions 2.3 and 2.4, we use

Lemma 2.6 Conditions 2.3 and 2.4 hold.

*Proof*: We begin by noting that

$$p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B)) = F(\frac{\theta_A + c_{\varepsilon}}{2})$$
  
=  $F(\frac{\theta_A + \theta_C^*}{2})$ 

where  $\theta_C^*$  is as defined in the proof of Lemma 2.5. So

$$\frac{\partial p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B))}{\partial \theta_A} = F'(\frac{\theta_A + \theta_C^*}{2})(1/2)(1 + \partial \theta_C^*/\partial \theta_A).$$

Differentiating the first-order condition gives  $\partial \theta_C^*/\partial \theta_A = 2F''(\frac{\theta_A + \theta_C^*}{2})/(F''(\frac{\bar{\theta}_B + \theta_C^*}{2}) - F''(\frac{\theta_A + \theta_C^*}{2}))$ . With the assumptions of Condition 2.3, both the denominator and numerator are negative. So  $\partial \theta_C^*/\partial \theta_A > 0$ . Thus,

$$\frac{\partial p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B))}{\partial \theta_A} = F'(\frac{\theta_A + \theta_C^*}{2})(1/2)(1 + \partial \theta_C^*/\partial \theta_A)$$
> 0.

This satisfies the first half of Condition 2.3. Next we look at  $p_A(\cdot, \bar{\theta}_B, l_{\varepsilon}(\theta_A, \bar{\theta}_B))$ .

From above, we know that  $l_{\varepsilon} = \theta_A - \varepsilon$ . So

$$p_A(\cdot,\bar{\theta}_B,l_{\varepsilon}(\theta_A,\bar{\theta}_B)) = F(\frac{\theta_A+\bar{\theta}_B}{2}) - F(\theta_A-\frac{\varepsilon}{2}).$$

Thus,

$$\frac{\partial p_A(\cdot, \bar{\theta}_B, l_{\varepsilon}(\theta_A, \bar{\theta}_B))}{\partial \theta_A} = (1/2)F'(\frac{\theta_A + \bar{\theta}_B}{2}) - F'(\theta_A - \frac{\varepsilon}{2}).$$

As  $\bar{\theta}_B$  is in some neighborhood of  $\theta_B^*$ , we have that the partial derivative described above is negative. So Condition 2.3 holds. A similar argument for candidate B shows that Condition 2.4 holds.

We now know that there is an equilibrium described by

$$L(\theta_A^*, \theta_B^*) = C(\theta_A^*, \theta_B^*) = R(\theta_A^*, \theta_B^*).$$

Of course, as Weber (1992), Theorems 2.6 and 2.8, has shown, this condition is equivalent to

$$F(\theta_A^*) = 1 - 2F(1/4 + \theta_A^*/2)$$
 and  $\theta_B^* = 1 - \theta_A^*$ .

Thus, the usual Palfrey equilibrium emerges as a special case of the general model we have developed.

#### Application to Multiple States

We are now ready to model some specific institutional features. Instead of representing voters by a density on [0,1], suppose we have groups of voters distributed over [0,1],  $[-\delta, 1-\delta]$ , and  $[\delta, 1+\delta]$  according to distributions  $F_1(x)$ ,  $F_2(x)$ , and  $F_3(x)$ . As

we are thinking of states in the Electoral College, we will denote the voters in the first interval as state 1, those in the second as state 2, and those in the last as state 3. In order to simplify the analysis, we assume that the voters in each interval have the same distribution F(x). Formally,

$$F_1(x) = F(x)$$

$$F_2(x) = F(x+\delta)$$

$$F_3(x) = F(x-\delta).$$

We also set  $\Theta = (\delta, 1 - \delta)$ . This restricts the policy positions of candidates to the (interior) overlap area of the three states. We discuss relaxing this restriction later in this section.

Although the states in American Electoral College system have varying weights, we simplify the analysis by requiring a candidate to receive a plurality in a simple majority of states in order to be elected. We assume that candidates maximize the probability of winning the election. This is the same as maximizing the probability of a plurality in at least two states, given the election rules specified above. Let the probability that candidate k receives a plurality of the votes in state i be denoted  $p_k^i$ . Then we have the following expression for the probability that a candidate wins the election:

**Lemma 2.7** The probability that candidate k wins the election is  $p_k^1 p_k^2 + p_k^1 p_k^3 + p_k^2 p_k^3 - 2p_k^1 p_k^2 p_k^3$ .

*Proof*:

$$\begin{split} P(\text{candidate }k \text{ wins election}) &= P(k \text{ wins } 1, 2, \text{ and } 3) + P(k \text{ wins } 1 \text{ and } 2) \\ &\quad + P(k \text{ wins } 1 \text{ and } 3) + P(k \text{ wins } 2 \text{ and } 3) \\ &= p_k^1 p_k^2 p_k^3 + p_k^1 p_k^2 (1 - p_k^3) + p_k^1 p_k^3 (1 - p_k^2) \\ &\quad + p_k^2 p_k^3 (1 - p_k^1) \\ &= p_k^1 p_k^2 + p_k^1 p_k^3 + p_k^2 p_k^3 - 2 p_k^1 p_k^2 p_k^3. \end{split}$$

Adopting the view that there is one voter in each state with an unknown ideal point, we suppose, as in the Palfrey model, that  $p_k^i = v_k^i$ , where  $v_k^i$  is the vote share that candidate k receives in a state i.

We have already derived  $v_k$  for state 1 (the unit interval). Similar expressions are obtained for states 2 and 3 by substituting  $F_k$  for F:

$$v_A^2(\theta_A, \theta_B, \theta_C) = \begin{cases} F_2(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ F_2(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_A < \theta_B < \theta_C \\ F_2(\frac{\theta_A + \theta_B}{2}) - F_2(\frac{\theta_C + \theta_A}{2}) & \text{if } \theta_C < \theta_A < \theta_B \\ \frac{1}{2}F_2(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A = \theta_B < \theta_C \\ \frac{1}{2}(1 - F_2(\frac{\theta_C + \theta_A}{2})) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$v_B^2(\theta_A, \theta_B, \theta_C) = \begin{cases} 1 - F_2(\frac{\theta_C + \theta_B}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F_2(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_C < \theta_A < \theta_B \\ F_2(\frac{\theta_B + \theta_C}{2}) - F_2(\frac{\theta_A + \theta_B}{2}) & \text{if } \theta_A < \theta_B < \theta_C \\ \frac{1}{2}F_2(\frac{\theta_B + \theta_C}{2}) & \text{if } \theta_A = \theta_B < \theta_C \\ \frac{1}{2}(1 - F_2(\frac{\theta_C + \theta_B}{2})) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$v_C^2(\theta_A, \theta_B, \theta_C) = \begin{cases} F_2(\frac{\theta_C + \theta_A}{2}) & \text{if } \theta_C < \theta_A \le \theta_B \\ F_2(\frac{\theta_C + \theta_B}{2}) - F_2(\frac{\theta_A + \theta_C}{2}) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F_2(\frac{\theta_B + \theta_C}{2}) & \text{if } \theta_A \le \theta_B < \theta_C. \end{cases}$$

Because  $F_2(x) = F(x + \delta)$ , we can rewrite these functions as

$$v_A^2(\theta_A, \theta_B, \theta_C) = \begin{cases} F(\frac{\theta_A + \theta_C}{2} + \delta) & \text{if } \theta_A < \theta_C < \theta_B \\ F(\frac{\theta_A + \theta_B}{2} + \delta) & \text{if } \theta_A < \theta_B < \theta_C \end{cases}$$

$$V_A^2(\theta_A, \theta_B, \theta_C) = \begin{cases} F(\frac{\theta_A + \theta_B}{2} + \delta) - F(\frac{\theta_C + \theta_A}{2} + \delta) & \text{if } \theta_C < \theta_A < \theta_B \\ \frac{1}{2}F(\frac{\theta_A + \theta_C}{2} + \delta) & \text{if } \theta_A = \theta_B < \theta_C \end{cases}$$

$$\frac{1}{2}(1 - F(\frac{\theta_C + \theta_A}{2} + \delta)) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$v_B^2(\theta_A, \theta_B, \theta_C) = \begin{cases} 1 - F(\frac{\theta_C + \theta_B}{2} + \delta) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F(\frac{\theta_A + \theta_B}{2} + \delta) & \text{if } \theta_C < \theta_A < \theta_B \\ F(\frac{\theta_B + \theta_C}{2} + \delta) - F(\frac{\theta_A + \theta_B}{2} + \delta) & \text{if } \theta_A < \theta_B < \theta_C \\ \frac{1}{2}F(\frac{\theta_B + \theta_C}{2} + \delta) & \text{if } \theta_A = \theta_B < \theta_C \\ \frac{1}{2}(1 - F(\frac{\theta_C + \theta_B}{2} + \delta)) & \text{if } \theta_C < \theta_A = \theta_B \end{cases}$$

$$v_C^2(\theta_A, \theta_B, \theta_C) = \begin{cases} F(\frac{\theta_C + \theta_A}{2} + \delta) & \text{if } \theta_C < \theta_A \le \theta_B \\ F(\frac{\theta_C + \theta_B}{2} + \delta) - F(\frac{\theta_A + \theta_C}{2} + \delta) & \text{if } \theta_A < \theta_C < \theta_B \\ 1 - F(\frac{\theta_B + \theta_C}{2} + \delta) & \text{if } \theta_A \le \theta_B < \theta_C. \end{cases}$$

Of course, similar expressions are obtained for state 3, by replacing  $+\delta$  with  $-\delta$ .

We will first concentrate on the entrant's probability of winning the election, given Lemma 2.7. It can be expressed as

$$p_{C}(\theta_{A}, \theta_{B}, \theta_{C}) = \begin{cases} F(\frac{\theta_{C} + \theta_{A}}{2})F(\frac{\theta_{C} + \theta_{A}}{2} + \delta) + F(\frac{\theta_{C} + \theta_{A}}{2}) \\ F(\frac{\theta_{C} + \theta_{A}}{2} - \delta) + F(\frac{\theta_{C} + \theta_{A}}{2} + \delta)F(\frac{\theta_{C} + \theta_{A}}{2} - \delta) \\ -2F(\frac{\theta_{C} + \theta_{A}}{2})F(\frac{\theta_{C} + \theta_{A}}{2} + \delta)F(\frac{\theta_{C} + \theta_{A}}{2} - \delta) \end{cases} & \text{if } \theta_{C} < \theta_{A} \le \theta_{B} \\ -2F(\frac{\theta_{C} + \theta_{B}}{2}) - F(\frac{\theta_{A} + \theta_{C}}{2}))(F(\frac{\theta_{C} + \theta_{B}}{2} + \delta) \\ -F(\frac{\theta_{A} + \theta_{C}}{2} + \delta)) + (F(\frac{\theta_{C} + \theta_{B}}{2}) - F(\frac{\theta_{A} + \theta_{C}}{2})) \\ (F(\frac{\theta_{C} + \theta_{B}}{2} - \delta) - F(\frac{\theta_{A} + \theta_{C}}{2} - \delta)) + (F(\frac{\theta_{C} + \theta_{B}}{2} + \delta) & \text{if } \theta_{A} < \theta_{C} < \theta_{B} \\ -F(\frac{\theta_{A} + \theta_{C}}{2} + \delta))(F(\frac{\theta_{C} + \theta_{B}}{2} - \delta) - F(\frac{\theta_{A} + \theta_{C}}{2} - \delta)) \\ -2(F(\frac{\theta_{C} + \theta_{B}}{2}) - F(\frac{\theta_{A} + \theta_{C}}{2}))(F(\frac{\theta_{C} + \theta_{B}}{2} + \delta) \\ -F(\frac{\theta_{A} + \theta_{C}}{2} + \delta))(F(\frac{\theta_{C} + \theta_{B}}{2} - \delta) - F(\frac{\theta_{A} + \theta_{C}}{2} - \delta)) \end{cases} \\ (1 - F(\frac{\theta_{B} + \theta_{C}}{2}))(1 - F(\frac{\theta_{B} + \theta_{C}}{2} - \delta)) \\ + (1 - F(\frac{\theta_{B} + \theta_{C}}{2} + \delta))(1 - F(\frac{\theta_{B} + \theta_{C}}{2} - \delta)) \\ -2(1 - F(\frac{\theta_{B} + \theta_{C}}{2}))(1 - F(\frac{\theta_{B} + \theta_{C}}{2} - \delta)) \\ -2(1 - F(\frac{\theta_{B} + \theta_{C}}{2}))(1 - F(\frac{\theta_{B} + \theta_{C}}{2} + \delta)) \\ (1 - F(\frac{\theta_{B} + \theta_{C}}{2} - \delta)). \end{cases}$$

We must check that this horrendous expression satisfies Assumptions 2.7 and 2.8 and Conditions 2.1 and 2.2. It is clear that Assumption 2.7 holds and careful inspection of the expression for  $p_C$  shows that Assumption 2.8 holds if F(x) is symmetric. We now turn our attention to Condition 2.2. In order to evaluate  $L(\theta_A, \theta_B)$ , consider

 $p_C$  when  $\theta_C \leq \theta_A - \varepsilon$ . Taking a partial with respect to  $\theta_A$  gives

$$\partial p_C/\partial \theta_A = \frac{1}{2} \{ F'(m_{AC}) [F(m_{AC} + \delta) + F(m_{AC} - \delta) - 2F(m_{AC} + \delta) F(m_{AC} - \delta) ]$$

$$+ F'(m_{AC} + \delta) [F(m_{AC}) + F(m_{AC} - \delta) - 2F(m_{AC}) F(m_{AC} - \delta) ]$$

$$+ F'(m_{AC} - \delta) [F(m_{AC} + \delta) + F(m_{AC}) - 2F(m_{AC} + \delta) F(m_{AC}) ] \}$$

where  $m_{AC}=(\theta_A+\theta_C)/2$ , the midpoint of candidate A and C's position. Applying the fact that F(x) is strictly increasing and thus F'(x)>0 and the fact that  $0 \le F(x) \le 1$  implies F(x)+F(y)-2F(x)F(y)>0 for  $x \ne y,^7$  we conclude that  $\partial p_C/\partial \theta_A>0$ . Thus, if the entrant must enter to the left of the established candidates, she will choose a position as close to them as possible. This implies that  $l_{\varepsilon}(\theta_A,\theta_B)=\theta_A-\varepsilon$  and  $l(\theta_A,\theta_B)=\theta_A$ . From this, and the expression for  $p_C$ , we can see that

$$L(\theta_A, \theta_B) = F(\theta_A)F(\theta_A + \delta) + F(\theta_A)F(\theta_A - \delta) + F(\theta_A + \delta)F(\theta_A - \delta) - 2F(\theta_A)F(\theta_A + \delta)F(\theta_A - \delta).$$

Taking a partial derivative with respect to  $\theta_A$  and proceeding as before shows us that  $L(\theta_A, \theta_B)$  is strictly increasing. Thus, part 1 of Condition 2.2 holds. Part 2 follows by symmetry and parts 3 through 6 are obviously satisfied. Thus, Condition 2.2 holds.

We now consider Condition 2.1. Lemma 2.5 says that Condition 2.1 holds for each state. By comparing the expressions for  $p_C$  in the single state model and the multiple

This is proved as follows. Since F(x) is increasing,  $(F(x) - F(y))^2 > 0$  for  $x \neq y$ . So  $F^2(x) + F^2(y) - 2F(x)F(y) > 0$ . But  $0 \leq F(x) \leq 1 \forall x$  implies  $F(x) \geq F^2(x) \forall x$ . Thus,  $F(x) + F(y) - 2F(x)F(y) \geq F(x) + F(y) - 2F(x)F(y) > 0$ .

state model, we can verify that

$$C(\theta_A, \theta_B) = C_1(\theta_A, \theta_B)C_2(\theta_A, \theta_B) + C_1(\theta_A, \theta_B)C_3(\theta_A, \theta_B)$$
$$+ C_2(\theta_A, \theta_B)C_3(\theta_A, \theta_B) - 2C_1(\theta_A, \theta_B)C_2(\theta_A, \theta_B)C_3(\theta_A, \theta_B)$$

where  $C_i(\theta_A, \theta_B)$  is the value of  $C(\theta_A, \theta_B)$  in state *i*. Fixing  $\bar{\theta}_B$ , we take a partial derivative with respect to  $\theta_A$ . We get:

$$\partial C(\theta_{A}, \theta_{B})/\partial \theta_{A} = C'_{1}(\theta_{A}, \theta_{B})[C_{2}(\theta_{A}, \theta_{B}) + C_{3}(\theta_{A}, \theta_{B}) - 2C_{2}(\theta_{A}, \theta_{B})C_{3}(\theta_{A}, \theta_{B})]$$

$$+ C'_{2}(\theta_{A}, \theta_{B})[C_{1}(\theta_{A}, \theta_{B}) + C_{3}(\theta_{A}, \theta_{B}) - 2C_{1}(\theta_{A}, \theta_{B})C_{3}(\theta_{A}, \theta_{B})]$$

$$+ C'_{3}(\theta_{A}, \theta_{B})[C_{1}(\theta_{A}, \theta_{B}) + C_{2}(\theta_{A}, \theta_{B}) - 2C_{1}(\theta_{A}, \theta_{B})C_{2}(\theta_{A}, \theta_{B})]$$

By Lemma 2.5,  $C'_i(\theta_A, \theta_B)$  is negative, and  $C_j + C_k - 2C_jC_k$  is positive because  $0 \le C_i \le 1$  and  $C_j \ne C_k$ . Therefore,  $\partial C(\theta_A, \theta_B)/\partial \theta_A$  is negative. So  $C(\cdot, \bar{\theta}_B)$  is continuous and strictly decreasing. In other words, part 1 of Condition 2.1 holds. A similar argument proves that part 2 holds. And clearly parts 3 and 4 hold. So we have confirmed both Conditions 2.1 and 2.2.

Now we can invoke Lemma 2.4 to show that there exists a unique pair  $(\theta_A^*, \theta_B^*)$  that satisfies

$$L(\theta_A^*, \theta_B^*) = C(\theta_A^*, \theta_B^*) = R(\theta_A^*, \theta_B^*).$$

To show that this pair is also an equilibrium, we must check Conditions 2.3 and 2.4. Lemma 2.6 shows that the conditions are true in each state. Once again, by comparing the expressions for  $p_A$  in the single state model and the multiple state

model, we can verify that

$$p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B)) = p_{A1}p_{A2} + p_{A1}p_{A3} + p_{A2}p_{A3} - 2p_{A1}p_{A2}p_{A3}$$

where  $p_{Ai} = p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B))$  in state *i*. Now we take a partial derivative with respect to  $\theta_A$ . We get

$$\begin{split} \partial p_A(\theta_A,\bar{\theta}_B,c_\varepsilon(\theta_A,\bar{\theta}_B))/\partial \theta_A &= p'_{A1}[p_{A2}+p_{A3}-2p_{A2}p_{A3}] \\ &+ p'_{A2}[p_{A1}+p_{A3}-2p_{A1}p_{A3}] + p'_{A3}[p_{A1}+p_{A2}-2p_{A1}p_{A2}]. \end{split}$$

By Lemma 2.6,  $p'_{Ai}$  is positive, and  $p_{Aj} + p_{Ak} - 2p_{Aj}p_{Ak}$  is positive because  $0 \le p_{Ai} \le 1$  and  $p_{Aj} \ne p_{Ak}$ . Therefore,  $\partial p_A/\partial \theta_A$  is positive. So  $p_A(\theta_A, \bar{\theta}_B, c_{\varepsilon}(\theta_A, \bar{\theta}_B))$  is continuous and strictly increasing. A similar argument proves that  $p_A(\cdot, \bar{\theta}_B, l_{\varepsilon}(\theta_A, \bar{\theta}_B))$  is strictly decreasing. So Condition 2.3 holds. By symmetry, the analogous argument confirms Condition 2.4.

By Theorem 2.3, therefore, the unique pair  $(\theta_A^*, \theta_B^*)$  that satisfies

$$L(\theta_A^*,\theta_B^*) = C(\theta_A^*,\theta_B^*) = R(\theta_A^*,\theta_B^*)$$

is a  $\varepsilon$ -distinguishable equilibrium.

### Equilibrium in Multiple States

Now that we know an equilibrium exists, we would like to understand more about it. Our first result is the following:

**Theorem 2.4** As  $\delta \to 0$ , the equilibrium of the multiple state model converges to the equilibrium of the single state model.

*Proof*: The equilibrium in the multiple state model is the pair  $(\theta_A^*, \theta_B^*)$  that satisfies

$$L(\theta_A^*, \theta_B^*) = C(\theta_A^*, \theta_B^*) = R(\theta_A^*, \theta_B^*).$$

From the expressions for  $L(\theta_A, \theta_B)$  and  $C(\theta_A, \theta_B)$  above, we see that

$$\lim_{\delta \to 0} L(\theta_A, \theta_B) = 3(L_1(\theta_A, \theta_B))^2 - 2(L_1(\theta_A, \theta_B))^3$$

and

$$\lim_{\delta \to 0} C(\theta_A, \theta_B) = 3(C_1(\theta_A, \theta_B))^2 - 2(C_1(\theta_A, \theta_B))^3,$$

where  $L_1$  and  $C_1$  are the L and C functions of state 1. But the equilibrium in the single state model satisfies

$$L_1(\theta_A, \theta_B) = C_1(\theta_A, \theta_B) = R_1(\theta_A, \theta_B),$$

so it must also satisfy

$$3(L_1(\theta_A, \theta_B))^2 - 2(L_1(\theta_A, \theta_B))^3 = 3(C_1(\theta_A, \theta_B))^2 - 2(C_1(\theta_A, \theta_B))^3.$$

This means that as  $\delta \to 0$ , the equilibrium of the multiple state model converges to the equilibrium of the single state model.

Thus, when the states converge to just a triplicate of the single state model, the equilibrium converges to the single state model's equilibrium. This is reassuring; the limiting case equilibrium is as it should be.

#### Example: Uniform Density of Voters

In order to illustrate the other interesting feature of the equilibrium, we assume that the density of voter ideal points is uniform. Formally, we suppose that F(x) = x. This allows us to explicitly solve out the equilibrium for a range of  $\delta$  values.

It should be noted, however, that this example illustrates some additional difficulties that arise in the case of increasing diversity of voter ideal points among the states. Recall that in the preceding analysis we assumed that the policy space available to the candidates,  $\Theta$ , was the (interior) overlap area of the three states,  $(\delta, 1 - \delta)$ . If the states have a significant divergence in voter ideal points, this overlap area could be small. It may be the case that a candidate may prefer to locate outside this overlap region. We allow for this possibility in the computations that follow. We also analyze the case of an empty overlap, which occurs if  $\delta > 1/2$ . If  $\delta \geq 1$ , then no two states intersect and it is obvious that if  $\theta_A = 0$  and  $\theta_B = 1$ , the entrant cannot win no matter where she chooses to enter.

It is clear that there is no equilibria in which  $\theta_A < 0.8$  So, for all  $\delta$ , we have  $\theta_A \ge 0$  and

$$\begin{split} L(\theta_A, \theta_B) &= L_1(\theta_A, \theta_B) L_2(\theta_A, \theta_B) + L_1(\theta_A, \theta_B) L_3(\theta_A, \theta_B) \\ &\quad + L_2(\theta_A, \theta_B) L_3(\theta_A, \theta_B) - 2 L_1(\theta_A, \theta_B) L_2(\theta_A, \theta_B) L_3(\theta_A, \theta_B) \\ &= \theta_A(\theta_A + \delta) + \theta_A(\max[\theta_A - \delta, 0]) \\ &\quad + (\theta_A + \delta)(\max[\theta_A - \delta, 0]) - 2\theta_A(\theta_A + \delta)(\max[\theta_A - \delta, 0]) \end{split}$$

<sup>&</sup>lt;sup>8</sup>For all  $\theta_A \leq 0$ ,  $L(\theta_A, \theta_B) = 0$ , and thus candidate A could safely gain by moving back toward the center.

and

$$C(\theta_A, \theta_B) = C_1(\theta_A, \theta_B)C_2(\theta_A, \theta_B) + C_1(\theta_A, \theta_B)C_3(\theta_A, \theta_B)$$
$$+ C_2(\theta_A, \theta_B)C_3(\theta_A, \theta_B) - 2C_1(\theta_A, \theta_B)C_2(\theta_A, \theta_B)C_3(\theta_A, \theta_B),$$

where

$$C_1(\theta_A, \theta_B) = \min\left[\frac{\theta_C^* + \theta_B}{2}, 1\right] - \max\left[\frac{\theta_A + \theta_C^*}{2}, 0\right]$$

$$C_2(\theta_A, \theta_B) = \max\left[\min\left[\frac{\theta_C^* + \theta_B}{2}, 1 + \delta\right] - \max\left[\frac{\theta_A + \theta_C^*}{2}, \delta\right], 0\right]$$

$$C_3(\theta_A, \theta_B) = \max\left[\min\left[\frac{\theta_C^* + \theta_B}{2}, 1 - \delta\right] - \max\left[\frac{\theta_A + \theta_C^*}{2}, -\delta\right], 0\right].$$

The optimal entry location,  $\theta_C^*$ , depends on the configuration of states, as determined by  $\delta$ . For  $\delta < 1/2$ , it is clear that by the symmetry of the equilibrium,  $c_{\varepsilon}(\theta_A, \theta_B) = 1/2 = c(\theta_A, \theta_B)$ . If there is no overlap among the three states, then locating at 1/2 may no longer be optimal for the entrant. For example, if  $\delta > 3/4$ , then, as  $\theta_A > 0$ , there is no chance that the entrant will win two or more states by locating at 1/2. In this case, the entrant should pick a location just inside one of the two established candidates' positions. In fact, for all  $\delta$  slightly larger than 1/2,  $c_{\varepsilon}(\theta_A, \theta_B) = \{\theta_A + \varepsilon, \theta_B - \varepsilon\}$  and thus  $\theta_C^* = c(\theta_A, \theta_B) = \{\theta_A, \theta_B\}$ . In addition, for all  $\delta$ , symmetry implies that, in equilibrium,  $\theta_A = 1 - \theta_B$ . By using these facts, we can solve the condition  $L(\theta_A, \theta_B) = C(\theta_A, \theta_B)$ .

For  $\delta < 1/2$ , we cannot explicitly solve this expression, but we can calculate a real solution which is graphed in Figure 2.1 on the following page. For  $\delta > 1/2$ , when

<sup>&</sup>lt;sup>9</sup>The exact value at which the entrant should switch to locating near one of the established candidates' positions is approximately  $\delta = .525$ .

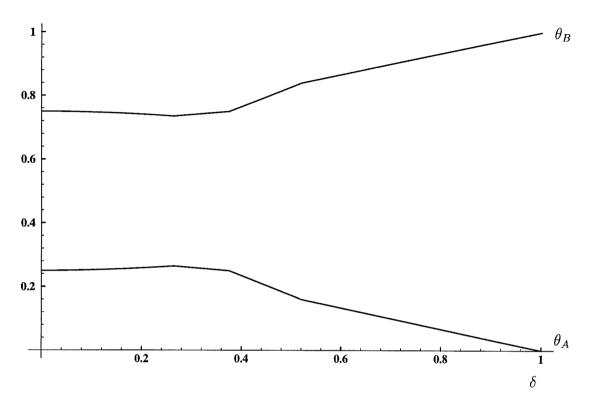


Figure 2.1: Established Candidate Positions and Electoral Diversity

 $\theta_C^* = \{\theta_A, \theta_B\}$ , we can solve this expression, as follows:

$$L(\theta_A, \theta_B) = C(\theta_A, \theta_B)$$
 
$$\theta_A(\theta_A + \delta) = (1/2 - \theta_A)(1 - \delta - \theta_A)$$
 
$$3\theta_A = 1 - \delta$$
 
$$\theta_A = (1 - \delta)/3$$

This solution is also incorporated into Figure 2.1. In the figure, the horizontal axis measures the diversity parameter  $\delta$  and the vertical axis graphs  $\theta_A$  and  $\theta_B$ , the candidates' equilibrium policy positions. As the figure illustrates, the equilibrium positions for the established candidates initially shift toward the center as the level of diversity

in the states increases, but then shift back towards the edges of the policy space with further increase of  $\delta$ . Specifically,  $\theta_A$  increases with  $\delta$  until a value  $\delta^*$  at which  $\theta_A(\delta^*) = \delta^* = 0.265$ . In this diversity region, with a uniform density, the probability remains constant in the center and the probability available in the extreme positions decrease. This motivates the two established candidates to shift toward the center as the threat of entry to the sides diminishes. Above this critical value, candidate position decreases as  $\delta$  continues to rise due to the shrinking vote share available to the center. Eventually, as  $\delta$  goes to 1, the established candidates locate at the extreme edges of the center state, preventing the entrant from having any chance of winning the election.

This is an important example because it indicates that some levels of diversity in the electorate can force the established candidates to adopt more, not less, similar positions. However, it should be noted that this analysis is based on the special case of the uniform density, and not on an arbitrary F(x). The extent to which the results differ for other distribution functions is an area that remains open.

### Interpretation of Voters

As we discussed previously, the standard approach admits two distinctly different interpretations with regard to the voters. The first view is that there are, in fact, an infinite number of voters distributed according to F(x). As an empirical matter, this means the entrant has a zero probability of winning the election. As a game-theoretic matter, this means that the candidates have perfect information about both the voter's location and preference. The second view is that there is a single voter with an unknown ideal point. This creates a positive probability that the entrant wins the election, and it introduces maximal uncertainty for the candidates. However,

the probability that an entrant wins may be too large in this interpretation. For the standard model with a uniform distribution, for example, the entrant has a one fourth probability of winning; a value higher than we may think is empirically supported.

In light of this, we would like to find an interpretation that predicts a positive, but small, probability of the entrant winning the election and also incorporates uncertainty into the candidates' decisions. We can achieve these aims by supposing that, instead of a single voter with an ideal point drawn from F(x), the model has a large, but finite, number of voters, each with an ideal point drawn from F(x). In this case, the candidates do not precisely know which voters will be casting ballots, and there is a small chance that the set of voters drawn will be clustered near the entrant's position and vote her into office.

Luckily, for simple one-dimensional models (such as Palfrey's entry model), maximizing the probability of winning the election is equivalent to maximizing the probability that a randomly selected voter will vote for you, which is equivalent to maximizing vote share in the traditional sense. Thus the equilibrium analysis will not change with this alternative view. However, in the multiple state models of the previous section, this equivalence does not persist. Indeed, it is easy to imagine situations in which a candidate may choose minimal support in one state in exchange for a larger probability of receiving a plurality in the other two states. This added strategic component greatly complicates the analysis, as we shall soon see.

In addition, such a view creates several theoretical challenges that must be dealt with. First of all, since the set of voters can be drawn in a multitude of different ways, it is difficult to calculate the probability that a given candidate wins the election. Indeed, with even the smallest of support, a candidate can find some selection of voters that will choose to elect him. Second, the probability of election can be highly variable. When the electorate is large, the probability of winning is much higher if

you have the sole largest support than if another candidate has a larger support than you. Thus, a small shift in a candidate's position can create large changes in the probability the candidate is elected. We will attempt to deal with these difficulties in the appendix.

Some results on approximation of the probability of winning are contained in the appendix. While much work remains to be done, we have derived an upper bound on the probability that a third-ranked candidate can win the election. This is presented in

**Lemma 2.8** If 
$$p_C^j < p_A^j$$
 and  $p_C^j < p_B^j$ , then  $p_C \le e^{-\frac{N}{2}(\sum_k \frac{1}{9p_k} - 1)}$ .

While this is an upper bound, the method of proof suggests that it may be legitimate to use 1/3 or  $\pi$  as an approximation of probability a third-ranked candidate wins the election. Clearly this is a fruitful area for further research.

### 2.4 Conclusion

The study of multicandidate elections is a crucial part of political science. While dual candidate contests have been modeled by the Median Voter Theorem, the question of third party entry has been largely ignored. Moreover, models that include institutional features in this context are indeed rare. This chapter has addressed several issues relating to modeling institutional features in multicandidate elections. First, we have shown the equivalence of three alternative equilibrium concepts. Second, we have developed a general framework for analyzing three candidate competition. We have applied this technique to a model motivated by the American presidential process and discovered that, in one case at least, increasing diversity in the electorate causes the established candidates to initially shift toward more moderate positions

and then back toward more extreme positions. Finally, prompted by our analysis of multiple state competition, we have suggested an alternative interpretation of the voters in the standard model. Thus, we have made progress in the area of modeling institutional features and have identified several areas that remain to be explored.

### **Appendix: Approximation Results**

This section addresses the mathematical details of the alternative view of voters presented above. We assume that the electorate consists of N voters with ideal points randomly distributed according to F(x). The candidates know the distribution F, but not the actual selection of ideal points for the N voters. In order to find the probability of a candidate k winning the election, the logical place to begin is the probability that a randomly selected voter will vote for k, given the candidates' positions. Since the voters are sincere, a random voter will vote for the candidate with position closest to the voter's ideal point. Thus, the probability a random voter votes for k is the probability that that voter's ideal point is closest to candidate k's position. But this is precisely what we termed candidate k's "vote share" in the first section. Thus, much of what we derived before will be applicable to the current problem.

So if we denote the probability that a randomly selected voter j will vote for a candidate k, given candidate positions  $\theta_A, \theta_B, \theta_C$ , as  $p_k^j(\theta_A, \theta_B, \theta_C)$ , we have that:

$$p_A^j(\theta_A, \theta_B, \theta_C) = v_A(\theta_A, \theta_B, \theta_C)$$
$$p_B^j(\theta_A, \theta_B, \theta_C) = v_B(\theta_A, \theta_B, \theta_C)$$
$$p_C^j(\theta_A, \theta_B, \theta_C) = v_C(\theta_A, \theta_B, \theta_C).$$

Also note that for fixed thetas,  $0 \le p_k^j \le 1$ , and  $p_A^j + p_B^j + p_C^j = 1$ . So for each of

the N voters in the electorate, the probability they cast their vote for candidate k is  $p_k^j$ . Probability theory tells us that in this situation the number of votes cast for each candidate can be described as a multinomial distribution or, more specifically, a trinomial distribution.<sup>10</sup> If we denote the number of votes that candidate k receives as  $n_k$ , then  $n_A + n_B + n_C = N$  and the likelihood is given by

$$P[N_A = n_A, N_B = n_B, N_C = n_C] = \frac{N!}{n_A! n_B! n_C!} (p_A^j)^{n_A} (p_B^j)^{n_B} (p_C^j)^{n_C}$$

where  $N_k$  is the random variable describing the number of votes cast for candidate k. With this expression, we can specify the probability that candidate k wins a plurality, which we denote  $p_k$ :<sup>11</sup>

$$p_k = \sum_{n_A, n_B, n_C: n_k > n_l, l \neq k} P[N_A = n_A, N_B = n_B, N_C = n_C]$$

Unfortunately, this expression is very difficult to evaluate and provides little guidance in linking candidate positions with probabilities of winning.

The most useful line of analysis in this regard is to approximate the trinomial with a more tractable expression. It is possible to achieve this with a  $\chi^2$  distribution.

**Lemma 2.9**  $P[N_A = n_A, N_B = n_B, N_C = n_C] \approx (2\pi N p_A^j p_B^j p_C^j)^{-\frac{1}{2}} \exp[-\frac{1}{2}X^2]$ , where  $X^2 = \sum_k \frac{(n_k - N p_k^i)^2}{N p_k^i}$  is approximately distributed as  $\chi^2$  with 2 degrees of freedom.

Proof: See Johnson and Kotz (1969), p. 285.

This may not seem to be much of an improvement; we have simply found an

<sup>&</sup>lt;sup>10</sup>See Dudewicz and Mishra (1988), p. 112

<sup>&</sup>lt;sup>11</sup>This is the same notation used to describe the probability that candidate k wins the election. In the standard model, winning the election is equivalent to winning a positive plurality, but in the multiple state model it is not. In the latter case, we will denote the probability that candidate k wins a plurality in state i as  $p_k^i$ .

alternative expression for the probability of a particular set of votes. But this expression can be used in a more creative way. Fix a particular distribution of votes to candidates,  $\bar{n}_A$ ,  $\bar{n}_B$ , and  $\bar{n}_C$ . Then vote outcomes that are less likely than this have higher values of  $X^2$ , and vote outcomes that are more likely have lower values of  $X^2$ . Therefore, this expression allows us to argue that the probability of getting an outcome that is equally or less likely than a given outcome is the probability that  $X^2$  is higher than the given value of  $X^2$ . This allows us to get a rough approximation of broad sets of vote distributions.

Notice that  $X^2$  has the form of a sum of square deviations. We actually must be concerned about two things. We must be satisfied that the  $\chi^2$  random variable is a good approximation of  $X^2$  and, second, that we have an accurate expression for the  $\chi^2$  random variable. Luckily, in our case, we have a closed form expression for the density of the  $\chi^2$  random variable.

**Lemma 2.10** A random variable Z with a  $\chi^2$  distribution with 2 degrees of freedom has density

$$f_Z(z) = \begin{cases} \frac{1}{2}e^{-z/2} & \text{if } 0 \le z < \infty \\ 0 & \text{otherwise} \end{cases}$$

*Proof*: Immediate from the definition of a  $\chi^2$  distribution. See Dudewicz and Mishra (1988).

**Lemma 2.11** If 
$$p_C^j < p_A^j$$
 and  $p_C^j < p_B^j$ , then  $p_C \le e^{-\frac{N}{2}(\sum_k \frac{1}{9p_k} - 1)}$ .

*Proof*: The most likely distribution of votes in which C wins the election is  $n_A = N/3 + 1$ ,  $n_B = N/3$  and  $n_C = N/3$ . Thus, the set of all vote distributions in which C wins the election must be a subset of the set of vote distributions that are equally

or less likely than this. That is

$$\begin{aligned} & \text{Prob}[\ C \ \text{wins}] \leq P[\chi^2 \geq \sum_k \frac{(N/3 - N p_k^i)^2}{N p_k^i}] \\ & = P[\chi^2 \geq N(\sum_k \frac{1}{9 p_k^i} - 1)] \\ & = \exp[-\frac{N}{2}(\sum_k \frac{1}{9 p_k^i} - 1)]. \end{aligned}$$

# Chapter 3

# Stability, Polls, and Duverger's

# Law

### 3.1 Introduction

How electoral laws affect the number and strength of political parties has been the focus of a great deal of work in political science. A centerpiece in this research effort is the famous work of Duverger (1954) on the effect of election type on political competition. There are two main statements to consider. The first, which Riker (1982) terms "Duverger's Law," is the proposition that "the simple-majority single-ballot [i.e., plurality] system favors the two-party system." (p. 217) In the next sentence, Duverger writes that "of all the hypotheses that have been defined in this book, this approaches the most nearly perhaps to a true sociological law." Indeed, Riker has defended Duverger's Law as an example of the cumulative progress of political science. The second proposition, termed "Duverger's Hypothesis" by Riker, states that "the simple-majority system with second ballot and proportional representation favor multi-partism." (p. 239)

We will focus attention on Duverger's Law, the more widely studied of Duverger's propositions. It has a rich intellectual history, which dates back to 1881 and an English barrister named Henry Droop who wrote

... the only explanation which seems to me to account for [two-party systems] is that the two opposing parties into which we find politicians divided in each of these countries [United Kingdom, United States, etc.] have been formed and are kept together by majority voting [what we now call plurality voting.] (quoted in Riker (1982), p. 756–757)

Since Droop's time, a broad range of research has touched on Duverger's Law. Several excellent reviews cover both the empirical testing of the Law and the theoretical explanations for its widespread applicability (Cox 1991; Duverger 1986; Riker 1982, 1986). It has garnered a good deal of empirical support, the best example of which remains the notable book by Rae (1967). From a theoretical perspective, Duverger offered the initial justification of his Law. He reasoned that the field of parties is trimmed to just two by the forces of "fusion" and elimination. Consider a third party (or candidate) receiving a small but significant number of votes. Facing a bleak electoral future, it may choose to ally itself with or even merge or "fuse" with one of the larger parties in exchange for consideration in other districts or a share of the benefits of office. This is the result of a strategic calculation made by the party leaders. If the third party does not go the route of fusion, it runs the risk that its voters will gradually desert it and the party will fall victim to elimination because of lack of support. This alternative is the result of strategic calculations by the voters.

<sup>&</sup>lt;sup>1</sup>As Riker (1982) notes: "It is customary to call the law by Duverger's name, not because he had much to do with developing it but rather he was the first to dare to claim it was a law. The memorial honors, therefore, a trait of character as much as a scientific breakthrough." (p. 754)

<sup>&</sup>lt;sup>2</sup>Duverger's use of the term "fusion" to describe mergers between political parties is sensible, but it runs counter to its common usage among American political scientists as "the joint sponsorship of a candidate by two or more parties." With this meaning, "fusion" is offered as an exception to Duverger's Law in which multiple parties persist in plurality rule districts. Examples are elections in the Midwest and West during the late nineteenth-century (Argersinger 1980) and contemporary New York (Cox 1987).

Duverger continues his reasoning for his Law: "Elimination in this sense (the second way in which bipartism is restored) is itself the result of two factors working together: a mechanical and a psychological factor." (p. 225-226) The mechanical effect refers to the exaggeration of the success of large parties in excess of their electoral margin of victory.<sup>3</sup> The psychological factor is the tendency of voters not to vote for parties that have no chance of winning. As Duverger puts it:

In cases where there are three parties operating under the simple-majority single-ballot system the electors soon realize that their votes are wasted if they continue to give them to the third party: whence their natural tendency to transfer their vote to the less evil of its two adversaries in order to prevent the success of the greater evil. This 'polarization' effect works to the detriment of a new party so long as it is the weakest party but it is turned against the less favoured of its older rivals as soon as the new party outstrips it. (p. 226)

This notion that unpopular parties do not receive votes because voters do not want to waste their votes is well known (Downs 1957; Schattschneider 1942). In 1869, Droop explained the concept this way:

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 or sets of 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), p. 756)

Droop's comments foreshadow the explanation of strategic voting offered by decision theory (Black 1978; Gutowski and Georges 1993; McKelvey and Ordeshook 1972). In this literature, an individual voter in isolation makes a voting decision based on exogenously determined probabilities of certain outcomes. If the voter's favorite candidate has a poor enough chance of winning, the voter may decide to vote

<sup>&</sup>lt;sup>3</sup>In American presidential election, the Electoral College has this effect. See Chapter 1.

for her more favored of the two viable candidates, thus avoiding a "wasted" vote. The logical next step is to ask if the exogenous probabilities can be made endogenous; is strategic voting supported even if voters account for the (possibly strategic) votes of others? Cox (1987a) was the first attempt to use this approach to analyze strategic voting in elections. By using the concept of Bayesian equilibrium (Harsanyi 1967-68), Cox was able to "close" the model; "by using an equilibrium concept one can take account of the logical connections that exist between preferences, beliefs about others' preferences, voting intentions, and beliefs about others' voting intentions." (p. 16) The paper's main contribution is to show that strategic voting is the norm rather than the exception and that this phenomenon tends to help a leading candidate and hurt a trailing candidate. Palfrey (1989) extends this approach to prove a much stronger result. Using a Bayesian game much like Cox's, he shows that as the size of an electorate grows without bound, the electoral support of all but the two leading candidates is completely eroded. Palfrey argues that, except for some knife-edge cases, "in equilibrium, the share of the vote for the third party declines to zero as the number of voters increases." Thus, Palfrey provides a strong theoretical argument in support of Duverger's Law.

There are several other equilibrium models of Duverger's Law that we will briefly mention.<sup>5</sup> Riker (1976) adds "disillusioned voting" to a model of sophisticated voting to explain the case of India. Wright and Riker (1989) compare plurality and runoff systems using theoretical models and empirical evidence. Feddersen (1992) proposes a model with several unique features. Unlike Cox and Palfrey, Feddersen examines a complete information game with costly voting. This cost of voting motivates voters

<sup>&</sup>lt;sup>4</sup>Ledyard (1984) first proposed such a "general equilibrium" model of voting, but he did not consider multiple candidates.

<sup>&</sup>lt;sup>5</sup>Much of this literature is surveyed by Cox (1987, 1991).

to avoid wasting their vote on a hopeless candidate. Thus, in equilibrium, only two two parties receive votes. Finally, Cox (1994) extends the Cox-Palfrey model to deal with multimember districts. He is able to prove the "M+1 rule" which states that strategic voting in M-seat districts has the effect that exactly M+1 candidates receive votes in equilibrium.

Myerson and Weber (1993) argue that Palfrey has overstated his case in "proving" Duverger's Law. They construct a quasi-Bayesian "voting equilibrium" in which all three candidates receive votes under plurality rule, in violation of Duverger's Law. They also show that this equilibrium is not just a knife-edge case, as claimed by Palfrey. They conclude that "Duverger's Law cannot be derived exclusively from analyses of voting equilibria." In addition, Myerson and Weber argue that "any derivation of Duverger's Law would seem to require some additional assumption of dynamic stability or persistence to eliminate equilibria of [this] type ... " In this chapter, we show that analyzing a dynamic process involving public opinion polls restores Duverger's Law by eliminating the "non-Duvergerian" equilibria.

The use of public opinion polls as a vehicle for information in the model links this work to the vast literature on public opinion polls, which we cannot hope to survey here. Indeed, polls have long been used as a tool to measure public opinion and other variables of interest to political scientists. But the opposite effect, in which poll results change voters' beliefs and thus their decisions, has received less attention in the literature. The possible "bandwagon" and "underdog" effects of election prediction via public opinion polls were first discussed by Simon (1954).<sup>6</sup> McKelvey and Ordeshook (1985a) offer a model in which straw polls serve to inform

<sup>&</sup>lt;sup>6</sup>Simon may have been the first scholar to analyze the possibility of "bandwagons," but ever since the first polls were taken people have recognized the possibility that their use could affect outcomes. See Smith (1990) and West (1991).

the electorate about the policy positions of candidates. Cukierman (1991) shows that pre-election polls can amplify the effects of shifts in ideology in the electorate.

In this chapter, we examine how public opinion polls can coordinate voters on particular Duvergerian equilibria and argue that coordination on a three-party equilibrium is impossible. The first section sets up the model and introduces the relevant concepts and notation for the chapter. The next section shows that equilibria exist in the model that violate Duverger's Law. The third section introduces a dynamic process involving a sequence of polls and shows that the non-Duvergerian equilibrium is unstable. The final section illustrates this instability and discusses how polls can act as a coordination device for voters.

#### 3.2 The Model

In this section, we develop the model we will use to examine the stability of non-Duvergerian equilibria in multi-candidate elections. By way of an example, Myerson and Weber (1993) have shown that such equilibria do exist and are immune to small changes in the parameters. The model described here is an extension of this example to the Palfrey model, in order to show that the same result holds. We can thus view this model as a hybrid of the models developed by Palfrey (1989) and Myerson and Weber (1993). The notation will follow that of Palfrey (1989).

There are three *candidates*, labeled A, B, and C and n voters. Each *voter* in the electorate has a preference ordering over the candidates that can be represented by a Von Neumann-Morgenstern utility function. These preferences are referred to as a voter's *type*. In the general model of Palfrey (1989), there are six groups of similar voter types that occur with positive probability, with each group of types correspond-

ing to one of the six possible (strict) preference orderings.<sup>7</sup> In addition, within each group, the utility of the second-ranked candidate is distributed according to some cumulative distribution function. In our model, we pick a particular specification of preferences suggested by an example examined by Myerson and Weber (1993).

We suppose that the electorate is composed of three groups of similar types. Two of these groups involve a uniform distribution over the utility of the second most preferred candidate, while the third is a "degenerate" group involving a single shared type representing one specific preference ranking. 8 So the electorate is described by three groups of voters and three probabilities that a randomly selected voter belongs to that group. The first group consists of voters who strictly prefer candidate C to the other two candidates, about whom they are indifferent. The probability a voter belongs to this group is denoted  $q_C$  which, following Myerson and Weber (1993), we assume is equal to .4. The second group consists of voters who (strictly) rank candidate A above candidate B in their orders, and who prefer either A or B to C. Specifically, we suppose that the utility values for the three candidates are normalized so that the most-preferred candidate, A, has utility 1, the least-preferred candidate, C, has utility 0, and voter  $\alpha$ 's utility for candidate B is  $v_{\alpha}$ . We refer to this group of voters as AB type voters, and a voter who assigns utility  $v \in (0,1)$  to B (the second-ranked candidate) as a AB-v type voter. We denote the probability that a randomly selected voter will be a AB type voter as  $q_{AB}$ . We assume the utility values v for candidate B for voters in this group are distributed uniformly on (0,1). The last group is made up of BA type voters, with utilities v also distributed uniformly on (0,1). That is, voters in this group have a (strict) preference ordering that ranks B

<sup>&</sup>lt;sup>7</sup>These orderings are ABC, ACB, BAC, BCA, CAB, and CBA.

<sup>&</sup>lt;sup>8</sup>Thus, we do not impose Palfrey's assumptions that "all preference types are possible" and that the "distribution of preferences has no mass points."

<sup>&</sup>lt;sup>9</sup>In the notation of Palfrey (1989), we are assuming that  $F_{AB}(\cdot)$  is uniform.

Voter	Utility for candidate		Prior	
Type	A	B	C	Probability
$\overline{AB}$	1	$v \sim U(0,1)$	0	$q_{AB} = .3 + \epsilon$
BA	$v \sim U(0,1)$	1	0	$q_{BA} = .3 - \epsilon$
C	0	0	1	$q_C = .4$

Table 3.1: Voter Types in the Electorate

above A above C and have utilities for A distributed uniformly on the unit interval. The probability that a randomly selected voter belongs to this group is denoted  $q_{BA}$ . In order to show the general existence of non-Duvergerian equilibria in this setting, we suppose that  $q_{AB} = .3 + \epsilon$  for some  $\epsilon > 0$ , and thus  $q_{BA} = .3 - \epsilon$ . We summarize the makeup of the electorate in Table 3.1. We denote the set of possible types by T. Formally, we define

$$T = (\{AB\} \times (0,1)) \bigcup (\{BA\} \times (0,1)) \bigcup (\{C\}).$$

The probability of a particular type  $t \in T$ , is given by a probability measure P which depends on  $q_{AB}$ ,  $q_{BA}$ , and  $q_C$ .

We have thus far laid out the distribution (in a probabilistic sense) of voters in the electorate. We now turn to formalizing plurality elections as a game among the voters. We suppose that each voter simultaneously decides which candidate to vote for, taking other voters' actions as given and maximizing expected utility. Voters receive the utility value of the candidate that is elected. In case of a tie, we assume that ties are broken alphabetically, i.e. A beats B in a tiebreaker. In Implicit in this description of the available actions is the assumption that all voters cast a ballot;

<sup>&</sup>lt;sup>10</sup>The results in this chapter do not rely on the particular choice of a tiebreaking rule. The rule we work with is computationally easiest.

voters are not permitted to abstain.

Rather than assume that each voter has complete information about the preferences of other voters, we suppose that voters possess incomplete information about the types of other voters. Voters have only (probabilistic) beliefs about the distribution of preferences in the electorate. Specifically, we assume that each voter knows their own preferences and believes that the preferences of the other voters are independently and identically distributed according to the true distribution described in the earlier paragraph. This is the standard assumption of common priors in Bayesian games.

In this Bayesian framework, we allow voters to choose actions that may depend on their type. Therefore, we describe voter i's strategies as (measurable) functions of the form  $\sigma^i: T \to \{A, B, C\}$ . These strategies represent the possible beliefs of other voters about what voter i will do, depending on i's type. In order to emphasize the anonymity of mass elections, we consider only symmetric Bayesian equilibrium. An equilibrium is symmetric if voters with identical types act identically. In other words, voters subscribe to a common function  $\sigma$  in equilibrium. This common, symmetric strategy is, in the language of Bayesian games,  $common\ knowledge$ .

In order to analyze the symmetric Bayesian equilibria in the game, we must characterize the best response of a voter if the other voters, distributed according to P, all use a strategy  $\sigma$ . Given the symmetric nature of the strategy, each voter faces an identical problem posed by the other n-1 voters, in the sense that the distribution of votes (of the other n-1 voters) is a repeated trinomial experiment, which is determined by  $\sigma$  and independent of i. Given a symmetric strategy  $\sigma$ , we can partition the type space T into sets  $D_A$ ,  $D_B$ , and  $D_C$ , such that  $D_k$  represents the types who

cast ballots for candidate k. Formally, we define

$$D_k = \{ t \in T : \sigma(t) = k \}.$$

In other words,  $D_k = \sigma^{-1}(k)$ . The probability that a randomly selected other voter will vote for candidate A, B, or C is denoted by  $\pi_A$ ,  $\pi_B$ , or  $\pi_C$ , respectively. These values are defined by

$$\pi_k = Pr[t \in T \mid \sigma(t) = k] = \int_{\sigma^{-1}(k)} dP(t) = \int_{D_k} dP(t).$$

Voters then use these probabilities to calculate the probability of any given distribution of votes. Specifically, the probability that candidate A gets (exactly) a votes, candidate B gets (exactly) b votes, and candidate C gets (exactly) c votes, out of the n-1 other voters is f(a,b,c), a trinomial probability with formula

$$f(x,y,z) = \frac{(n-1)!}{x!y!z!} \pi_A^x \pi_B^y \pi_C^z, \tag{3.1}$$

where x, y and z are non-negative integers satisfying x + y + z = n - 1.

In addition to restricting our attention to symmetric equilibria, we also rule out any equilibria that involves voters using weakly dominated strategies. Voting for a particular candidate is weakly dominated if voting for some other candidate would sometimes be better for the voter, and never worse. For example, an AB-v voter should never cast a ballot for C; some of the time a vote for A might make a difference in electing A, but never will it make the voter worse off than choosing candidate C. In fact, the following is easily shown to be true:

Lemma 3.1 In any equilibria that does not involve weakly dominated strategies, all

of the type C voters cast ballots for C, and the other two types of voters never vote for candidate C.

Thus, type C voters always vote sincerely in the model. This lemma implies that in any equilibrium,  $\pi_C = q_C = .4$ .

Although the C type voters always vote sincerely, the other voters may choose to vote for their second choice in order to avoid "wasting" their vote. The conditions under which such strategic voting will occur are presented in the following lemma which is due to Palfrey (1989).

**Lemma 3.2** In an equilibrium involving probabilities  $\pi_A$ ,  $\pi_B$ , and  $\pi_C$ , if voter i is an AB-v type voter, then i's best response is:

- 1. Vote for A if  $(1-v)p_{AB}^{n} + p_{AC}^{n} > vp_{CB}^{n}$ .
- 2. Vote for B if  $(1-v)p_{AB}^n + p_{AC}^n < vp_{CB}^n$ .
- 3. Either vote for A or B if equality holds.

If voter i is an BA-v type voter, then i's best response is:

- 1. Vote for B if  $(1 v)p_{AB}^n + p_{CB}^n > vp_{AC}^n$ .
- 2. Vote for A if  $(1 v)p_{AB}^n + p_{CB}^n < vp_{AC}^n$ .
- 3. Either vote for B or A if equality holds,

where

 $p_{AB}^{n}$  = probability that voting for A yields A, but voting for B yields B.

 $p_{AC}^n$  = probability that voting for A yields A, but voting for B yields C.

 $p_{CB}^n$  = probability that voting for A yields C, but voting for B yields B.

The  $p_{ij}^n$  probabilities are essentially the probabilities that candidates i and j will tie for first place.<sup>11</sup> They depend on the probabilities that a randomly sampled other voter will vote for a particular candidate,  $(\pi_A, \pi_B, \text{ and } \pi_C)$ , and the size of the electorate, n. In the following trinomial formulas, we use notation such as  $\{b < a = c - 1\}_n$  for the set of non-negative integers (a, b, c) satisfying b < a = c - 1 and a + b + c = n - 1. We thus have:

$$\begin{split} p_{AC}^n &= \text{ probability of } \{b < a = c - 1\}_n \\ &= \sum_{\{b < a = c - 1\}_n} f(a, b, c) \\ &= \sum_{k = \left[\frac{n-1}{3}\right] - 1} \frac{(n-1)!}{k!(k+1)!(n-2k-2)!} (\pi_A)^k (\pi_B)^{n-2k-2} (\pi_C)^{k+1}, \end{split}$$

where [z] is the smallest integer not less than z.

$$\begin{split} p_{CB}^n &= \text{ probability of } \{a < b = c - 1\}_n \\ &= \sum_{\{a < b = c - 1\}_n} f(a, b, c) \\ &= \sum_{k = \left[\frac{n-1}{3}\right] - 1} \frac{(n-1)!}{k!(k+1)!(n-2k-2)!} (\pi_A)^{n-2k-2} (\pi_B)^k (\pi_C)^{k+1}. \end{split}$$

We note for later use the fact that  $p_{CB}^n$  can be expressed as  $\sum_{\{b < a = c-1\}_n} f(b, a, c)$ .

The Because of the way ties are broken, the vote totals used in calculating the  $p_{ij}^n$  probabilities are more complicated. For example,  $p_{AC}^n$  is the probability that a voters vote for A, b voters vote for B, and c voters vote for C, out of the other n-1 voters, where b < a and a = c-1.

Also,

$$\begin{split} p_{AB}^n &= \text{ probability of } \{a=b \geq c-1\}_n + \text{ probability of } \{a=b-1 \geq c-1\}_n \\ &= \sum_{\{a=b \geq c-1\}_n} f(a,b,c) + \sum_{\{a=b-1 \geq c-1\}_n} f(a,b,c) \\ &= \sum_{k=\left[\frac{n}{2}\right]-1}^{\left[\frac{n}{2}\right]-1} \frac{(n-1)!}{k!k!(n-2k-1)!} (\pi_A)^k (\pi_B)^k (\pi_C)^{n-2k-1} \\ &+ \sum_{k=\left[\frac{n}{3}\right]-1}^{\left[\frac{n-1}{2}\right]-1} \frac{(n-1)!}{k!(k+1)!(n-2k-2)!} (\pi_A)^k (\pi_B)^{k+1} (\pi_C)^{n-2k-2}. \end{split}$$

Now, if  $\sigma$  is an equilibrium, it must be the case that all voters are using a best response to  $\sigma$ . As Palfrey points out, this implies that an equilibrium can be characterized by two values,  $v_{AB}^*$  and  $v_{BA}^*$ , known as *cutpoints*. These cutpoints represent the (maximum) level of utility that a voter must have in order to vote sincerely. Thus, if a voter has utility  $v > v_{ij}^*$  for her second choice, then that voter will vote strategically for her second choice, candidate j. Note that if  $v_{ij}^* = 1$ , then all type ij voters will vote sincerely in equilibrium. It is sufficient to characterize equilibria by cutpoints for only type AB and BA voters, as type C voters must use the weakly dominant strategy of voting for C, i.e.  $v_C^* = 1$ .

Therefore the following condition must hold in equilibrium:

Condition 3.1 In an equilibrium involving probabilities  $\pi_A$ ,  $\pi_B$ , and  $\pi_C$ , type C voters vote for C, and type AB and BA voters use the strategy "Vote for i if  $v < v_{ij}^*$  and vote for j if  $v > v_{ij}^*$ ," where

$$v_{ij} = \min \left\{ 1, \frac{p_{ij}^n + p_{ik}^n}{p_{ij}^n + p_{kj}^n} \right\}.$$

However, this is only a partial description of an equilibrium. We add to this a rational

expectations condition common to Bayesian equilibria. Specifically, we require that in a given equilibrium, the  $\pi$ 's that are taken as given in the first condition are in fact generated by the distribution of voters and the strategies they use in that equilibrium. In this way, voters' beliefs about the actions of other voters (the  $\pi$ 's) are self-fulfilling expectations, i.e. the beliefs are not contradicted in equilibrium. Given the special distributions of voter preferences we have defined, this rational expectations condition simplifies to

Condition 3.2 In an equilibrium characterized by cutpoints  $v_{AB}^n$  and  $v_{BA}^n$ , the following holds:

$$\pi_A = q_{AB}v_{AB}^n + q_{BA}(1 - v_{BA}^n) = (.3 + \epsilon)v_{AB}^n + (.3 - \epsilon)(1 - v_{BA}^n),$$

$$\pi_B = q_{BA}v_{BA}^n + q_{AB}(1 - v_{AB}^n) = (.3 - \epsilon)v_{BA}^n + (.3 + \epsilon)(1 - v_{AB}^n),$$

and  $\pi_C = q_C v_C^n = .4$ .

So we have described the two conditions that must be met in an equilibrium in this model. It is natural to ask what such equilibria look like, if they exist. Palfrey (1989) develops a (partial) answer to this question. First, any single party equilibrium, in which only one candidate receives votes, must involve some voters employing weakly dominated strategies. Thus we rule out any such equilibria in the model. Second, Palfrey shows that multiple two-party equilibria exist. There are two such equilibria in the model, one in which candidates A and C receive all the votes cast and one in which only candidates B and C receive votes. We conclude this section by stating Palfrey's theorem dealing with these two-party equilibria. In the next section, we discuss the existence of three-party equilibria in the model. Finally, in a later section we address the problem of coordinating on one of the two possible two-party equilibria.

It turns out the key element in the voter's decision calculus is the relative likelihood of a tie between the leading and second-place candidate versus a tie between the trailing candidate and one of the other candidates. Specifically, Palfrey proves the following lemma:

**Lemma 3.3 (Palfrey)** Assume 
$$\pi_A < \min\{\pi_B, \pi_C\}$$
. Then  $\lim_{n\to\infty} \left(\frac{p_{AC}^n}{p_{CB}^n}\right) = 0$  and  $\lim_{n\to\infty} \left(\frac{p_{AB}^n}{p_{CB}^n}\right) = 0$ .

This lemma is used to prove the following theorem, applicable to our model:

**Theorem 3.1 (Palfrey)** There does not exist a sequence of equilibria  $\{v_n^*\}_{n=1}^{\infty}$  such that  $0 < \pi_A^* < \min\{\pi_B^*, \pi_C^*\}$ , where  $\pi_j^* = \lim_{n \to \infty} \pi_j(v_n^*)$ , j = A, B, C and  $\pi_j(v_n^*)$  is given by Condition 2.

This theorem states that a trailing candidate's vote share must erode away to nothing as the size of the electorate gets large, if all voters vote strategically.

## 3.3 Existence of Non-Duvergerian Equilibria

In the previous section, we developed a model that admitted two equilibria supportive of Duverger's Law. In these equilibria, the "wasted vote" phenomena led to some voters voting strategically for their second choice in the election. As a result, only two candidates received votes in the limit of the equilibria. In this section, we examine the possible existence of equilibria that do not support Duverger's Law. We will show that for a range of  $\epsilon$ , there exist three-party equilibria and that these equilibria persist as the number of voters grows without bound. The latter statement is important because the two-party equilibria identified in the previous section are truly two-party only in the limit. In any large, finite electorate, i.e.  $n < \infty$ , these equilibria still contain a small number of voters voting for the trailing candidate. This follows from

the fact that in any finite electorate the probability of a tie between the last-place candidate and the leading candidate is still positive, albeit very small. Thus, voters who prefer the trailing candidate with sufficiently small utility for the second-place candidate will still vote sincerely.

As suggested by Myerson and Weber (1993), we will examine the situation described as an "exceptional case" by Palfrey, namely the case in which  $\pi_A^* = \pi_B^* < \pi_C^*$ . In this situation, one dominant candidate emerges, opposed by two smaller parties. As the two smaller parties receive (nearly) equal vote shares, supporters of neither are willing to abandon their most preferred candidate; there is no single trailing candidate for voters to abandon, ala Duverger's Law.

In the non-Duvergerian equilibrium we identify, equality of vote shares,  $\pi_A^* = \pi_B^*$ , holds only in the limit as n goes to infinity. In a finite electorate, the following happens in equilibrium: the more popular candidate, A, receives a slightly smaller vote share than B. This prompts the small fraction of AB voters with very high utilities for candidate B to vote strategically for B. Thus, even though there is a larger distribution of AB type voters than BA types, the lower vote share for A is self-fulfilling. As the size of the electorate grows, the vote share for candidate A increases while the vote share for B decreases. In fact, the shares for both candidates approach A. In the limit, the candidates receive equal vote shares and thus a three-party equilibrium is sustained. This argument is laid out formally in the rest of this section.

We will make use of a lemma first proved by Cox (1987a).

Lemma 3.4 (Cox) Suppose  $\pi_A > \pi_B$ . Then

$$f(a,b,c) > f(b,a,c) \iff a > b,$$

where f(x, y, z) is defined by equation 3.1.

*Proof*: From the definition of f, we have

$$f(a,b,c) > f(b,a,c) \Longleftrightarrow \pi_A^a \pi_B^b > \pi_A^b \pi_B^a \Longleftrightarrow \pi_A^{(a-b)} > \pi_B^{(a-b)}$$
.

As  $\pi_A > \pi_B$ ,

$$\pi_A^{(a-b)} > \pi_B^{(a-b)} \iff a-b > 0 \iff a > b.$$

In order to show the existence of a non-Duvergerian equilibrium in the model, we must identify values  $v_{AB}^n$  and  $v_{BA}^n$  that satisfy the conditions for equilibria given in the previous section. One element of the preceding discussion is that these values must vary for different values of n. Luckily, we do not have to consider every combination of  $v_{AB}^n$  and  $v_{BA}^n$  for every possible n. To show this, we establish two lemmas. They both involve a critical value  $\tilde{v} = .3/(.3 + \epsilon)$ . If  $v_{BA}^n = 1$ , then  $\tilde{v}$  is the value of  $v_{AB}^n$  which insures that the probability of a vote for candidate A and candidate B are equal, i.e.  $\pi_A = \pi_B$ . The first lemma deals with the possibility of equilibria involving AB cutpoints greater than  $\tilde{v}$ .

**Lemma 3.5** In equilibrium, it cannot be the case that  $\tilde{v} < v_{AB}^n < 1$ .

Proof: If  $\tilde{v} < v_{AB}^n < 1$ , then by Condition 2,  $\pi_A > .3$ . Thus, it must be that  $\pi_A > \pi_B$ . By Lemma 3.4, if a > b, then f(a, b, c) > f(b, a, c). Therefore,

$$\sum_{\{b < a = c - 1\}_n} f(a, b, c) > \sum_{\{b < a = c - 1\}_n} f(b, a, c),$$

<sup>&</sup>lt;sup>12</sup>This follows immediately from Condition 2.

and thus  $p_{AC}^n > p_{CB}^n$ . However, by Condition 1,  $v_{AB}^n$  must satisfy

$$v_{AB}^{n} = \min \left\{ 1, \frac{p_{AB}^{n} + p_{AC}^{n}}{p_{AB}^{n} + p_{CB}^{n}} \right\}.$$

Thus  $v_{AB}^n = 1$  in equilibrium, which is a contradiction.

This lemma does not rule out an equilibrium with  $v_{AB}^n = 1$ . Such an equilibrium would have  $\pi_A > \pi_B$  and  $\pi_C > \pi_B$  for all n. Thus, this equilibrium is one of the Duvergerian equilibria accounted for by Palfrey in Theorem 3.1. In order to identify a non-Duvergerian equilibrium we must focus on equilibria with  $v_{AB}^n < \tilde{v}$ . The following lemma shows that any equilibria in this region must involve BA type voters voting sincerely.

**Lemma 3.6** If  $v_{AB}^n < \tilde{v}$  holds in some equilibrium, then  $v_{BA}^n = 1$ .

*Proof*: Follows immediately from Condition 1 and the supposition that  $v_{AB}^n < 1$ .

This lemma allows us to ignore  $v_{BA}^n$  in what follows and focus on the cutpoint for AB type voters,  $v_{AB}^n$ , which we denote as v. As an additional convenience, we define a "best response" function  $h_n(v)$  that summarizes the two conditions for equilibria. Specifically, if a value  $v_{AB}^n = v < \tilde{v}$  is an equilibrium, Condition 2 defines  $\pi_A(v) = (.3 + \epsilon)v$  as  $v_{BA}^n = 1$ . It follows that  $\pi_B(v) = .6 - \pi_A(v)$ , as  $\pi_C = .4$ . Thus, by Condition 1, we define the "best response" to v as

$$h_n(v) = \min \left\{ 1, \frac{p_{AB}(\pi_A(v)) + p_{AC}(\pi_A(v))}{p_{AB}(\pi_A(v)) + p_{CB}(\pi_A(v))} \right\}.$$

It is immediately apparent that  $h_n(v)$  is continuous,  $h_n(v) = 1$  for  $v > \tilde{v}$ , and  $h_n(v)$  is differentiable everywhere except  $\tilde{v}$ . It follows from the definition that v is an equilibrium if and only if v is a fixed point of  $h_n(v)$ , i.e.  $h_n(v) = v$ . For any n, v = 0

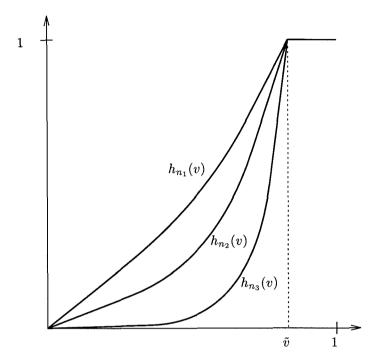


Figure 3.1:  $h_n(v)$  for three values of n

is a fixed point of  $h_n(v)$ . This corresponds to the two-party equilibria in which all voters abandon candidate A. We now proceed to show that there is also a three-party equilibria in the model. We shall use the following lemma:

**Lemma 3.7** For any fixed  $v < \tilde{v}$ ,  $\lim_{n \to \infty} h_n(v) = 0$ .

*Proof*: We can rewrite the definition of  $h_n(v)$  as

$$h_n(v) = \min \left\{ 1, \frac{\frac{p_{AB}(\pi_A(v))}{p_{CB}(\pi_A(v))} + \frac{p_{AC}(\pi_A(v))}{p_{CB}(\pi_A(v))}}{\frac{p_{AB}(\pi_A(v))}{p_{CB}(\pi_A(v))} + 1} \right\}.$$

From Lemma 3.3 it follows that  $\lim_{n\to\infty} h_n(v) = 0$ .

We illustrate the implication of these facts about  $h_n(v)$  in Figure 3.1. The figure graphs  $h_n(v)$  for three different values of n,  $n_1 < n_2 < n_3$ . As illustrated, the function moves closer to the horizontal axis and the dotted line at  $\tilde{v}$  as n grows larger.

We next prove the main result of this section on the existence of a three-party equilibria in the model. This theorem establishes that non-Duvergerian equilibria are possible under a range of preference distributions. Thus, the equilibria in the model are not generically Duvergerian.

**Theorem 3.2** There exists a sequence of equilibria  $\{v_n^*\}_{n=1}^{\infty}$  such that  $0 < \pi_A^* = \pi_B^* < \pi_C^*$ , where  $\pi_j^* = \lim_{n \to \infty} \pi_j(v_n^*)$ , j = A, B, C and  $\pi_j(v_n^*)$  is given by Condition 2.

Proof: For each n, let  $F_n$  be the set of fixed points of  $h_n(v)$ . As  $0 \in F_n$  for all n,  $F_n$  is non-empty. Fix  $v < \tilde{v}$  and  $\delta < 0$ . Lemma 3.7 is equivalent to the statement that there exists an integer N such that for all  $n \ge N$ ,  $h_n(v) < \delta$ . As  $h_n$  is continuous and  $h_n(v) = 1 > v$  for  $v > \tilde{v}$ , this implies that for sufficiently large n, each  $F_n$  contains a fixed point between v and  $\tilde{v}$ . More specifically, for any  $v < \tilde{v}$ , there exists an integer  $N_v$  such that  $\forall n \ge N_v$ ,  $(v, \tilde{v}) \cap F_n \ne \emptyset$ . Thus it is possible to choose a sequence of fixed points that converge to  $\tilde{v}$ . Denote this convergent sequence as  $\{v_n^*\}_{n=1}^{\infty}$ . This sequence fulfills the requirements of the theorem.

This theorem is illustrated in Figure 3.2 on the following page. For each n, there is an equilibrium  $v_n^*$  satisfying  $v_n^* = h_n(v_n^*)$ . We show three such values in the figure. The figure also shows that as n gets large, the shape of the  $h_n$  function forces  $v_n^*$  to converge to  $\tilde{v}$ .

## 3.4 Stability and Dynamics of Non-Duvergerian Equilibria

In the last section, we demonstrated the existence of non-Duvergerian, three-party equilibria in the model. This equilibrium is susceptible to two types of criticisms.

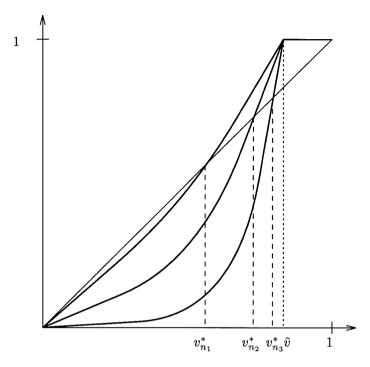


Figure 3.2: Sequence of Equilibria Converging to  $\tilde{v}$ 

The first critique is that it requires widespread, accurate information that is unlikely to be present in a real electorate. The share of AB voters who vote strategically for candidate B in equilibrium depends critically on the size of the electorate, n, and the exact proportion of AB voters in the population,  $\epsilon$ . The larger n is, the more AB type voters that are "allowed" by the equilibrium to vote sincerely. While we have assumed that these values are common knowledge among the voters, it is highly doubtful that such an assumption would be justified in a real electorate. This feature is particularly striking as the other, two-party equilibria of the model do not depend critically on these values. Even with some disagreement or uncertainty among the electorate about the parameters, either of the two-party equilibrium could be sustained. These equilibria require only that voters agree about which candidate is trailing. In the non-Duvergerian equilibrium, however, uncertainty could cause

some voters to vote "incorrectly" (i.e. not as prescribed by the equilibrium). Will the political process restore the delicate balance required by the equilibrium? This question raises the issue of *stability* of an equilibrium, which is a major focus of this section.

The other criticism of the three-party equilibrium in the model also applies to the Duvergerian equilibria. As Cox (1991) points out, all of the equilibria we have identified require voters to overcome a very serious coordination problem. The "fulfilled expectations" assumption acts to sidestep this issue. For example, the non-Duvergerian equilibrium requires that just the right number of AB types vote for their second choice, even though they outnumber the BA type voters; a number that insures candidate A trails B by just enough to justify the strategic voting. Likewise, the Duvergerian equilibria, in which either candidate A or B gets no votes, requires that voters all know, without communication, which candidate is the viable candidate that will receive all of the votes. Cox likens this issue to the coordination problem posed by the Battle of the Sexes game and comments that "these models should both be interpreted as showing that, if coordination problems can be overcome, then all equilibria in multicandidate plurality elections entail that only two candidates will receive any support." (emphasis as in original, p. 13) While this qualifier is certainly necessary, our analysis will have a great deal to say about overcoming the problem of coordination in multicandidate elections.

The root cause of both of these weaknesses lies in the use of a static model to describe a dynamic process. As Palfrey puts it, "the results in this essay have been interpreted as implying that if three parties are competing in a single-ballot, winner-take-all system, eventually (i.e., in an equilibrium) one of the parties will be weeded out. However, the model is static, even though the interpretation has dynamic overtones." (emphasis added, p. 85) Indeed, the reasoning Duverger offers for his Law is

a dynamic story in which voters, over time, gradually abandon an unpopular party in larger and larger numbers until no support remains. The static "one-shot" noncooperative game we have used to model the main elements of this story cannot fully describe the dynamics involved.

Both of these objections arise because voters in the model acquire information, either about the parameters of the electorate or the viability of candidates, without an explanation for how this (self-fulfilling) information arises in the population. Both Palfrey and Myerson and Weber appeal to public opinion polls, past election outcomes, and media reports, among others, as sources of the equilibrium information of voters. These polls, election outcomes, etc. may depend on earlier polls which may depend on even earlier polls, and so forth. With this interpretation, an equilibrium occurs when the most recent poll provides the electorate with enough information to act in a way that justifies the predictions of the poll. However, neither paper explicitly models the dynamic process of pre-election polling or considers the stability of the resulting equilibrium. It is this task to which we now turn.

In the analysis, we will use the function  $h_n(v)$  defined in the previous section to describe an equilibrium in terms of fixed points. In this section, we think of  $h_n(v)$  as describing an adjustment process that gives a best response for an individual AB type voter to the shared belief about the current cutpoint used by the AB types as a group.<sup>13</sup> This description is based on the following (tatonnement) process. Let  $v_0 < \tilde{v}$  be some initial cutpoint used by AB type voters. Suppose everyone begins by using  $v_0$  in making their vote decisions. Next suppose that a public opinion poll is taken to measure voting intention and the results are publicly announced. Then, given

 $<sup>^{13}</sup>$ As was shown in the previous section, as long as  $v < \tilde{v}$  the best response for BA type voters and C type voters is to vote sincerely. Thus we once again focus attention on AB type voters and suppose that the other voter types all vote sincerely.

the cutpoints in use, the results should be, on average, that candidate A receives a fraction  $q_{AB}v_0$  of the vote, candidate B gets  $q_{BA} + q_{AB}(1 - v_0)$  of the vote, and C receives support from  $q_C = .4$  of the voters. Suppose voters interpret these results as the probabilities that a randomly selected other voter will vote for a particular candidate, which we have called  $\pi_A$ ,  $\pi_B$ , and  $\pi_C$ . Then the best response under these expectations is to use a cutpoint  $v_1$  satisfying:

$$v_1 = h_n(v_0) = \min \left\{ 1, \frac{p_{AB}(\pi_A(v_0)) + p_{AC}(\pi_A(v_0))}{p_{AB}(\pi_A(v_0)) + p_{CB}(\pi_A(v_0))} \right\}.$$

Now suppose that there is a second opinion poll. Using the cutpoint  $v_1$ , voters' responses once again are announced and voters use these results to update their beliefs about the probabilities of votes from other voters. This gives rise to a new cutpoint  $v_2 = h_n(v_1)$ . If this process continues with a large number of polls, we can evaluate the convergence properties of the cutpoints used by the voters, where these cutpoints are given by the sequence  $v_{t+1} = h_n(v_t)$ .

This approach can also be used to determine the stability of the various equilibria in the model. Suppose that voters initially use cutpoints that differ slightly from those prescribed by the equilibrium. If the polling process just described leads voters to adjust their behavior so as to be more in accordance with the equilibrium, we say the equilibrium is *stable*. If the polling process causes voters to use cutpoints that differ even more widely from the equilibrium cutpoints, we say the equilibrium is *unstable*.

This concept of stability of an equilibrium is captured by a definition of stability first developed by Palfrey and Rosenthal (1991) for a similar problem. Their definition, in the context of our model, is the following:

#### **Definition 3.1** An equilibrium $v^*$ is:

- an expectationally stable equilibrium (ESE) if there exists an interval  $I(v^*)$  containing  $v^*$  such that, for all  $v_0 \in I(v^*)$ ,  $[v_0 h_n(v_0)](v_0 v^*)] > 0$  if  $v_0 \neq v^*$ .
- a globally expectationally stable equilibrium (GESE) if it is an ESE relative to the open interval  $(0, \tilde{v})$ .
- an expectationally unstable equilibrium (EUE) if there exists an interval  $I(v^*)$  containing  $v^*$  such that, for all  $v_0 \in I(v^*)$ ,  $[v_0 h_n(v_0)](v_0 v^*)] < 0$  if  $v_0 \neq v^*$ .

Note that our definition of an EUE is a stronger statement about the dynamic process than  $v^*$  is not an ESE. An EUE involves a dynamic process which moves cutpoints further away from the equilibrium cutpoints, as opposed to a process which does not move the cutpoints closer.

In order to evaluate the stability of our equilibrium, we need the following lemma:

**Lemma 3.8** There exists an integer N such that for all n > N,  $\frac{d}{dv}h_n(v) > 1$  at  $v_n^*$ , where  $v_n^*$  is an element of the sequence of equilibria described by Theorem 3.2.

*Proof*: Differentiating  $h_n$  yields

$$\frac{d}{dv}h_n(v) = \frac{(p_{AB} + p_{CB})\frac{d}{dv}(p_{AB} + p_{AC}) - (p_{AB} + p_{AC})\frac{d}{dv}(p_{AB} + p_{CB})}{(p_{AB} + p_{CB})^2}.$$

Since  $h_n(v_n^*) = v_n^*$  at an equilibrium  $v_n^*$ , we can rewrite this expression as

$$\frac{d}{dv}h_n(v_n^*) = \frac{\frac{d}{dv}(p_{AB} + p_{AC})(v_n^*) - v_n^* \frac{d}{dv}(p_{AB} + p_{CB})(v_n^*)}{p_{AB}(v_n^*) + p_{CB}(v_n^*)}.$$

Thus we need the derivatives of each of the  $p_{ij}$ 's.

$$\begin{split} \frac{d}{dv}p_{AC}^n &= \frac{d}{dv}\pi_A(v)\frac{d}{d\pi_A}p_{AC}^n \\ &= (q_{AB})\left(\sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1}f(k,n-2k-2,k+1)\left[\frac{k}{\pi_A}-\frac{n-2k-2}{\pi_B}\right]\right) \\ &= q_{AB}p_{AC}^n\sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1}\frac{f(k,n-2k-2,k+1)}{p_{AC}^n}\left[k\left(\frac{1}{\pi_A}+\frac{2}{\pi_B}\right)-\frac{(n-2)}{\pi_B}\right] \\ &= q_{AB}p_{AC}^n\left[\left(\frac{1}{\pi_A}+\frac{2}{\pi_B}\right)\sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1}k\frac{f(k,n-2k-2,k+1)}{p_{AC}^n} \\ &-\frac{(n-2)}{\pi_B}\cdot\frac{\sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1}f(k,n-2k-2,k+1)}{p_{AC}^n}\right] \\ &= q_{AB}p_{AC}^n\left[\left(\frac{1}{\pi_A}+\frac{2}{\pi_B}\right)(nE_{AC}^n)-\frac{(n-2)}{\pi_B}(1)\right], \end{split}$$

where  $E_{AC}^n = \frac{1}{n} \sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1} k \frac{f(k,n-2k-2,k+1)}{p_{AC}^n}$ . Likewise we can show that

$$\frac{d}{dv}p_{CB}^{n} = q_{AB}p_{CB}^{n} \left[ \frac{(n-2)}{\pi_{A}} - \left( \frac{1}{\pi_{B}} + \frac{2}{\pi_{A}} \right) (nE_{BC}^{n}) \right],$$

where  $E_{BC}^n = \frac{1}{n} \sum_{k=\left[\frac{n-1}{3}\right]}^{\left[\frac{n-1}{2}\right]-1} k \frac{f(n-2k-2,k,k+1)}{p_{CB}^n}$  and

$$\frac{d}{dv}p_{AB}^{n} = q_{AB} \left[ p_{AB}^{n} \left( \frac{1}{\pi_{A}} - \frac{1}{\pi_{B}} \right) (nE_{AB}^{n}) - \frac{\sum_{k=\left[\frac{n}{3}\right]-1}^{\left[\frac{n-1}{2}\right]-1} f(k,k+1,n-2k-2)}{\pi_{B}} \right],$$

where  $E_{AB}^{n} = \frac{1}{n} \left( \sum_{k=\left[\frac{n+1}{3}\right]-1}^{\left[\frac{n}{2}\right]-1} k \frac{f(k,k,n-2k-1)}{p_{AB}^{n}} + \sum_{k=\left[\frac{n}{3}\right]-1}^{\left[\frac{n-1}{2}\right]-1} k \frac{f(k,k+1,n-2k-2)}{p_{AB}^{n}} \right)$ . The  $E_{ij}^{n}$  terms represent the expected winning shares of the two candidates who receive the most

votes, conditional on there being a tie for first place between candidates i and j, and Palfrey (1989) has shown that these terms have the following limits:

$$\lim_{n \to \infty} E_{AC}^n = \max \left\{ \frac{1}{3}, \frac{1}{2 + \frac{\pi_B}{\sqrt{\pi_A \pi_C}}} \right\}$$

$$\lim_{n \to \infty} E_{CB}^n = \frac{1}{2 + \frac{\pi_A}{\sqrt{\pi_B \pi_C}}}$$

$$\lim_{n \to \infty} E_{AB}^n = \frac{1}{3}.$$

For sufficiently large n,  $v_n^*$  is close to  $\tilde{v}$  and thus  $\lim_{n\to\infty} E_{AC}^n = 1/(2 + \frac{\pi_B}{\sqrt{\pi_A \pi_C}})$ . Returning to  $h_n(v)$  and plugging in the derivatives gives:

$$\begin{split} \frac{d}{dv}h_n(v_n^*) &= q_{AB}n \, \frac{p_{AB}^n \left(\frac{1}{\pi_A} - \frac{1}{\pi_B}\right) E_{AB}^n + p_{AC}^n \left[\left(\frac{1}{\pi_A} + \frac{2}{\pi_B}\right) E_{AC}^n - \frac{1}{\pi_B}\right]}{p_{AB}(v_n^*) + p_{CB}(v_n^*)} \\ &- v_n^* q_{AB}n \, \frac{p_{AB}^n \left(\frac{1}{\pi_A} - \frac{1}{\pi_B}\right) E_{AB}^n + p_{CB}^n \left[\frac{1}{\pi_A} - \left(\frac{1}{\pi_B} + \frac{2}{\pi_A}\right) E_{CB}^n\right]}{p_{AB}(v_n^*) + p_{CB}(v_n^*)} \\ &+ q_{AB} \, \frac{\left(1 - v_n^*\right) \frac{\sum_{k=\lfloor \frac{n}{3} \rfloor - 1}^{\lfloor \frac{n-1}{2} \rfloor - 1} f(k, k+1, n-2k-2)}{\pi_B} + \frac{2p_{AC}^n}{\pi_B} + \frac{2p_{CB}^n}{\pi_B}}{p_{AB}(v_n^*) + p_{CB}(v_n^*)}} \\ &= q_{AB}n \, \frac{\frac{p_{AB}^n}{p_{CB}} \left(\frac{1}{\pi_A} - \frac{1}{\pi_B}\right) E_{AB}^n + \frac{p_{AC}^n}{p_{CB}} \left[\left(\frac{1}{\pi_A} + \frac{2}{\pi_B}\right) E_{AC}^n - \frac{1}{\pi_B}\right]}{(p_{AB}(v_n^*)/p_{CB}(v_n^*)) + 1}} \\ &- v_n^* q_{AB}n \, \frac{\frac{p_{AB}^n}{p_{CB}} \left(\frac{1}{\pi_A} - \frac{1}{\pi_B}\right) E_{AB}^n + (1) \left[\frac{1}{\pi_A} - \left(\frac{1}{\pi_B} + \frac{2}{\pi_A}\right) E_{CB}^n\right]}{(p_{AB}(v_n^*)/p_{CB}(v_n^*)) + 1} \\ &+ \frac{O(1)}{(p_{AB}(v_n^*)/p_{CB}(v_n^*)) + 1}, \end{split}$$

where we use the notation O(1) to stand for terms that are bounded as n grows without bound.

 $<sup>^{-14}</sup>$ Actually, because of the tie-breaking rule, this is not exactly true. For example,  $E_{AC}^n$  is the expected vote share of A conditional on B coming in last and A receiving exactly one less vote than C.

Now, applying Lemma 3.3 and using the fact that all of the  $E_{ij}^n$  terms have constant limits, we get that

$$\begin{split} \frac{d}{dv}h_n(v_n^*) &= q_{AB}n \, \frac{(o(1) \cdot O(1)) + (o(1) \cdot O(1))}{o(1) + 1} \\ &- \pi_A(v_n^*)n \, \frac{(o(1) \cdot O(1)) + \left[\frac{1}{\pi_A} - \left(\frac{1}{\pi_B} + \frac{2}{\pi_A}\right) E_{CB}^n\right]}{o(1) + 1} + \frac{O(1)}{o(1) + 1} \\ &= \frac{n\left(o(1) - \left[1 - \left(\frac{\pi_A}{\pi_B} + 2\right) E_{CB}^n\right]\right) + O(1)}{o(1) + 1}, \end{split}$$

where we use the notation o(1) to stand for terms that go to 0 as n grows without bound.

For sufficiently large n, we have

$$\frac{d}{dv}h_n(v_n^*) \sim \frac{n\left[\left(2 + \frac{\pi_A}{\pi_B}\right)\left(\frac{1}{2 + \frac{\pi_A}{\sqrt{\pi_B \pi_C}}}\right) - 1 + o(1)\right] + O(1)}{o(1) + 1}.$$

As  $2 + \frac{\pi_A}{\pi_B} > 2 + \frac{\pi_A}{\sqrt{\pi_B \pi_C}}$  when  $\pi_B < \pi_C$  (true for sufficiently large n), we conclude that  $\lim_{n \to \infty} \frac{d}{dv} h_n(v_n^*) = \infty$ . Thus there exists an integer N such that for all n > N,  $\frac{d}{dv} h_n(v) > 1$  at  $v_n^*$ .

While all that we need for our theorem is that  $h'_n(v) > 1$ , it should be noted that we actually prove that  $h'_n(v)$  is of order n. This should not be surprising, as we know that for any  $v < \tilde{v}$ ,  $h_n(v)$  is near 0 for sufficiently large n while  $h_n(v) = 1$  for  $v > \tilde{v}$ .

The main result of this section is that the non-Duvergerian equilibrium in the model is expectationally unstable.

**Theorem 3.3** For sufficiently large n, the equilibria  $v_n^*$  described by Theorem 3.2 are expectationally unstable.

*Proof*: We must show that there exists an integer N such that for all n > N,  $v_n^*$  is an EUE. By Lemma 3.8, such an N exists such that  $\frac{d}{dv}h_n(v_n^*) > 1$ . This implies that there exists an interval  $I(v_n^*)$  containing  $v^*$  such that, for all  $v \in I(v_n^*)$ ,

$$\frac{h_n(v) - h_n(v_n^*)}{v - v_n^*} > 1.$$

As  $v_n^*$  is an equilibrium,  $h_n(v_n^*) = v_n^*$ . Thus

$$\frac{h_n(v) - v_n^*}{v - v_n^*} = \frac{h_n(v) - v + v - v_n^*}{v - v_n^*} > 1 \Longrightarrow \frac{h_n(v) - v}{v - v_n^*} > 0.$$

We conclude that

$$\frac{v - h_n(v)}{v - v_n^*} < 0.$$

Thus, if  $v \neq v_n^*$ ,  $v - h_n(v)$  and  $v - v_n^*$  have opposite sign. So for all n > N,  $(v - h_n(v))(v - v_n^*) > 0$  for all  $v \in I(v_n^*)$  such that  $v \neq v_n^*$ . Thus  $v_n^*$  is an EUE.

Therefore, if voters are not perfectly informed as to the "correct" equilibrium cutpoint, even by a minute amount, the polling process will move the cutpoints used by the voters away from the three-candidate equilibrium.

### 3.5 Polls as a Coordination Device

In the last section, we showed that the non-Duvergerian equilibrium in the model is expectationally unstable. In fact, we can prove more than this. In this section, we show that the equilibrium is globally unstable and the dynamic polling process must converge to one of the two Duvergerian equilibria, depending on the initial beliefs of voters. Thus, the polling process acts as a coordination device for the voters.

The key step in justifying these conclusions is the following lemma:

**Lemma 3.9** For sufficiently large n,  $h_n(v) < v$  for  $0 < v < v_n^*$  and  $h_n(v) > v$  for  $v > v_n^*$ .

Proof: The second half of the lemma is trivial. For the first half, suppose the result does not hold. Then for some sequence  $\{v_n\}_{n=N}^{\infty}$ ,  $h_n(v_n) > v_n$  for  $0 < v_n < v_n^*$ . As  $h_n(v)$  is continuous, this implies that for all n > N, there exists a fixed point of  $h_n(v)$  at v > 0. Assume this sequence of fixed points converges, otherwise consider a convergent subsequence. If this sequence converges to  $\tilde{v}$ , this contradicts Theorem 3.3 (which applies to any sequence of fixed points of  $h_n(v)$  converging to  $\tilde{v}$ ). If it does not, it contradicts Theorem 3.1, as no such three-party equilibrium can persist.

This lemma implies that the "best response" function  $h_n(v)$  has a form like that of Figure 3.3. Moreover, it implies that the non-Duvergerian equilibrium is globally unstable and that the dynamic polling process converges to one of the two Duvergerian equilibria. This is also illustrated in Figure 3.3. Suppose voters hold an initial belief  $v_0$  that is less than the equilibrium cutpoint  $v_n^*$ . Then the lemma implies that the adjusted cutpoint  $v_1 = h_n(v_0)$  will be less than  $v_0$ . Likewise,  $v_2$  will be less than  $v_1$  and, as the figure illustrates, eventually  $v_t$  must converge to 0. A similar argument shows that if  $v_0$  is above the equilibrium cutpoint,  $v_t$  converges to 1. Thus, the repeated cycles of opinion poll and adjustment lead voters to one of the Duvergerian equilibria.

In fact, this dynamic process can shed some light on overcoming the coordination problem faced by voters. Polls can guide voters to coordinate on a particular equilibrium, depending on their initial beliefs. In the previous section, the voters shared an initial belief  $v_0$  about the cutpoint used by the AB type voters, which was then adjusted to  $v_1$  after the first poll, and so on. This is a strong assumption about

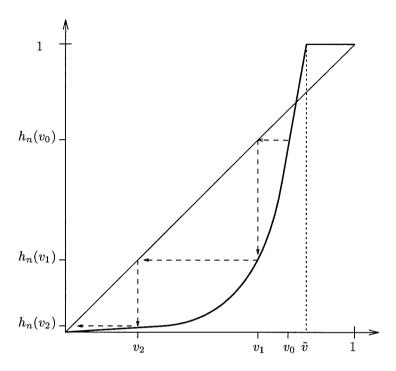


Figure 3.3:  $h_n(v)$  and the Adjustment Path of Opinion Polls

the initial conditions of the polling process. Instead, we relax this assumption and suppose that each voter  $\alpha$  has some initial belief  $v_0^{\alpha}$  about  $v_{AB}$ .<sup>15</sup> When the first poll is taken, suppose each voter votes according to the initial belief  $v_0^{\alpha}$ . This poll will result in outcomes  $s_A$ ,  $s_B$  and  $s_C$  for the candidates, where these shares depend on the distribution of beliefs. Suppose that after this first poll, voters adjust their beliefs that these shares represent the probabilities that a randomly selected other voter will vote for a given candidate. Then the polling process serves to aggregate the intentions of the voters and transmit this information about the viability of the candidates back to the electorate. If the polling reveals that candidate B is the serious challenger to C, then a bandwagon effect develops as more and more AB type voters switch their vote to A. Likewise, if A has the strongest initial support, a bandwagon develops for

<sup>&</sup>lt;sup>15</sup>For completeness, we should add an initial belief about  $v_{BA}$ .

A at the expense of candidate B.

In this way, polls can help voters coordinate on the viable candidates in the race, to insure that their votes are not wasted. However, this coordination can be fully effective only after an infinite number of polls. Thus, in actual elections we should not expect to see such dramatic results. In fact, if polls are not frequent enough or are ignored by voters, this process may have little effect on the actual outcome. But as polling becomes an even bigger part of media coverage of elections, we should expect that strategic voting and coordination effects should have an ever greater importance in politics.

### 3.6 Conclusion

In this chapter, we have focused on the "wasted vote" phenomenon and Duverger's Law. We have shown that in a Bayesian game model of strategic voting, there do exist non-Duvergerian equilibria in which all three candidates receive votes (in the limit). However, these equilibria require extreme coordination and any variation in beliefs leads voters away from them to one of the Duvergerian equilibria. Thus, non-Duvergerian equilibria are unstable, while two-party equilibria are not.

In addition, we describe how pre-election polls provide information to voters about the viability of candidates and can thus be used by voters to coordinate on a Duvergerian outcome. This conclusion about the behavior of voters should be testable. While there are no empirical studies that directly address this issue, there has been experimental confirmation of this prediction. Forsythe et al. (1993) describe a set of plurality rule experiments that closely resemble the theoretical model developed here. They show that pre-election polls significantly reduce the frequency with which the Condorcet loser is elected. With polls, voters are able to coordinate on one the

Duvergerian equilibria. While these results are promising, more work needs to be done on this process in the lab and among real electorates and elections.

## Chapter 4

## Information, Polls, and

## Bandwagons

#### 4.1 Introduction

A central concept in the classic theory of democracy and representative government is the importance of a well informed electorate. Citizens, the argument goes, must have sufficient knowledge of the issues and choices present in the political system if they are to delegate and monitor the actions of their representatives. Indeed, political philosophers have long seen public political awareness and sophistication as essential requirements for a successful democratic system. Early empirical work on voter surveys dealt a critical blow to the foundation of this view (Berelson et al. 1954; Campbell et al. 1960; Converse 1964). The classic image of the democratic citizen was found to be far removed from the reality of the average voter. Campbell et al. (1960) documented an appallingly low level of ideological awareness and political understanding by the American electorate. The authors describe their landmark book *The American Voter* as a "portrait of an electorate almost wholly without

detailed information about decision making in government." (p. 543) The impact of these findings on classic democratic theory was addressed by Berelson et al. (1954): "our data... reveal that certain requirements commonly assumed for the successful operation of democracy are not met by the behavior of the 'average' citizen." (p. 307)

These early findings sparked an intense debate in the field of electoral research over the level of political knowledge in the population on both substantive and methodological grounds (Nie and Andersen 1974; Achen 1975; Nie et al. 1976; Converse and Markus 1979; Smith 1989; excellent reviews are Niemi and Weisberg 1993, ch. 4; Dalton and Wattenberg 1993). While this issue has not been fully resolved, it is safe to say that modern electorates do not match the lofty ideals of classic democratic theory.

This empirical conclusion also has bearing on the numerous rational choice models of voter decision making that assume complete information on the part of voters and candidates, such as the celebrated Median Voter Theorem (Black 1958; Downs 1957; Enelow and Hinich 1984). Critics argue that as these models assume levels of information and sophistication that are unmet by contemporary electorates, their implications are vacuous.<sup>1</sup>

In both the empirical and theoretical studies of this issue, a central question that arises is how do voters make political choices based on their limited political knowledge. One approach to this question, that dates back to Downs, is to analyze the cues and political shortcuts used by voters to overcome their informational deficiencies. Downs lists four sources of information that voters could use to supplement their

<sup>&</sup>lt;sup>1</sup>Indeed, the fact that Downs, one the founders of positive political theory, devoted an entire chapter in his groundbreaking 1957 book to shortcuts used by voters to reduce their information costs appears to have been overlooked by many rational choice authors.

own limited knowledge, political parties, newspapers and television, interest groups, and other voters — friends, family, or neighbors. In his book *The Reasoning Voter*, Popkin (1991) argues that "low-information rationality" best describes the decision making process used by "reasoning voters." The term "low-information rationality" refers to "the information shortcuts and rules of thumb that voters use to obtain and evaluate information and to choose among candidates." In Popkin's view, the use of such shortcuts is unavoidable — "an inescapable fact of life [that] will occur no matter how educated we are, how much information we have, and how much thinking we do." (p. 218)

McKelvey and Ordeshook formalized this intuition that voters without perfect information can acquire useful information from other sources. In a series of papers (McKelvey and Ordeshook 1985a, 1985b, 1986b), they develop formal models of elections in which uninformed voters are able to gather valuable information, through contemporaneous poll results or past election results, and, in equilibrium, vote as if they had full information. Candidates in these models may also be afflicted with limited information, but once again enough information is transmitted by polls or past elections to insure that both candidates converge to the median voter's ideal point, just as in a complete information setting. In other papers, the authors extend this line of research to include multidimensional settings (McKelvey and Ordeshook 1984, 1987), influential interest groups (McKelvey and Ordeshook 1986a), and experimental evidence in support of their findings (McKelvey and Ordeshook 1990a, 1990b).

Some authors argue that the very nature of modern democratic systems insure the persistence of low levels of information in the electorate (Downs 1957, Popkin et al. 1976). If acquiring political information is a costly activity, voters must weigh this cost against the meager gain of an informed vote making a difference in the unlikely event that that vote changes the outcome of an election. The authors of *The* 

American Voter, Campbell et al. (1960), put it best:

[A] part of the electorate is no doubt capable of informing itself about political matters but is unwilling to pay the cost that such information-getting would entail. Very few people seem motivated strongly enough to obtain the information needed to develop a sensitive understanding of decision making in government. It is a rather unusual individual whose deeper personality needs are engaged by politics, and in terms of rational self-interest, the stakes do not seem to be great enough for the ordinary citizen to justify his expending the effort necessary to make himself well informed politically. (pp. 543-544)

This argument is intensified by the results of McKelvey and Ordeshook indicating that enough information is available from low-cost sources like public opinion polls to insure a correct vote. Thus, a voter may rationally choose to "free-ride" rather than invest in political information gathering.

Our purpose in this chapter is not to delve deeply into the countervailing normative implications offered by these theories of decision making by voters. Rather, we limit ourselves to analyzing the relationship between voter information and election outcomes in a multicandidate setting. Voters, with limited information, may find it even more difficult to make a correct choice when faced with more than two candidates. In a two candidate race, if a voter dislikes candidate A, a reasonable course of action, especially for a voter with limited information, would be to vote for A's opponent. In a multicandidate race, however, a voter who dislikes candidate A is unable to apply this simple shortcut; the voter is still faced with choosing between several alternative candidates.

In this chapter, we address the role that limited information may play in multicandidate elections. We extend a model of two candidate elections originally developed by McKelvey and Ordeshook to the multicandidate case and obtain several results on the ability of uninformed voters to gather and apply information available to them to their voting decisions. In the next section, we define the relevant components of the model and describe the information structure in the electorate. In section 4.3 we examine "voter equilibria" in which voters vote optimally, given their beliefs, and these beliefs are self-fulfilling in the sense that they are not contradicted by observable information. We present two results in this framework. First, there is an essentially unique voter equilibrium conditional on a fixed profile of candidate positions and voter beliefs, in which all voters, informed and uninformed alike, vote as if they had perfect information. Second, we define a dynamic process involving a sequence of straw votes such that this fully revealing equilibrium is attained regardless of the initial conditions of the process. The following section extends this analysis to include uncertainty among candidates, as well as voters. Once again, enough information is transmitted in equilibrium so that the candidates act as if they were fully informed. Specifically, we show that for three candidate competition, no equilibrium exists in a standard model of position taking, but an essentially unique equilibrium exists for a model of third party entry. Finally, in section 4.5 we show that if a small minority of voters are fully informed and use this information to vote strategically, in equilibrium all voters, including uninformed sincere voters, act as if they were voting strategically based on full information. Thus, the "wasted vote" phenomenon can be viewed as the result of voter's low-information rationality.

Finally, it should be noted that although we use the terminology of polls and straw votes in the construction of the model, we intend these terms to be interpreted broadly. In the formal model developed in this chapter, voters observe the results of straw polls and use this information to adjust their beliefs about the candidates. However, this information need not be transmitted in the form of a public opinion poll. Voters are certainly exposed to a great deal of "horse-race" coverage in election campaigns, but there are also other means of acquiring this information such as

general media exposure of candidates, discussions with friends and co-workers, and newspaper and interest group endorsements. Thus, the theory developed here should be seen as a general description of information transmission in the electorate and not a narrow model of election polling and politics.

#### 4.2 Notation and Definitions

In this section we introduce the notation and definitions used in developing the model in this chapter. We will follow the notation used in McKelvey and Ordeshook (1985a). We begin by describing the *voters* in the model by a probability measure space  $(\Omega, \mathcal{F}, \mu)$ . The set  $\Omega$  is partitioned into two subsets, U and I, which represent the uninformed and informed voters in the model, respectively. The set of alternatives is a closed convex set  $X \subseteq \mathbb{R}$ . The ideal points of voters are assigned by a  $(\mathcal{F}$  measurable) function  $y^* : \Omega \to X$ . Thus, voter  $\alpha$ 's ideal point is given by  $y^*(\alpha)$ , which is also denoted by  $y^*_{\alpha}$ . We construct the cumulative distribution function of the voters' ideal points in the usual way.<sup>2</sup> For any (measurable) set  $C \subseteq \Omega$  and  $x \in \mathbb{R}$ , define  $F_C(x) = \mu(\{\alpha \in C \mid y^*_{\alpha} \le x\})$ . We use the shorthand notation  $F(x) = F_{\Omega}(x)$ . In addition, each voter has a utility function given by  $u_{\alpha} : X \to \mathbb{R}$ , where  $u_{\alpha}(x) = u(x, y^*_{\alpha})$  and u(x, y) is a function that is symmetric and single-peaked in x about y, for all  $y \in X$ . Thus, voter  $\alpha$ 's utility for an alternative  $x \in X$  is  $u_{\alpha}(x)$ , a function that is symmetric and single-peaked about  $y^*(\alpha)$ .

The k candidates in the model are given by the set  $K = \{1, 2, ..., k\}$ . With each candidate we associate a point  $s_j \in X$ , candidate j's policy position or platform. We denote the platform profile of all of the candidates as  $s = (s_1, s_2, ..., s_k)$ . Without loss

<sup>&</sup>lt;sup>2</sup>See Chung (1974) for a development of this subject. Also note that we do not require that a density function exist for the distribution.

of generality, we assume that the candidates are labeled such that  $s_1 \leq s_2 \leq \cdots \leq s_k$ . For later use, we denote the set of functions from K into X as S, with typical element s(j).

In the model, the voters can each cast a vote for exactly one candidate; they may not abstain. Thus the action set of voter  $\alpha$  is K, the set of candidates. The set of possible ballots that the voters as a whole could cast is given by B, the set of  $(\mathcal{F})$  measurable functions from  $\Omega$  into K. Elements of B are given by  $b:\Omega\to K$ . Thus, voter  $\alpha$ 's ballot is given by  $b(\alpha)$ , which is also denoted by  $b_{\alpha}$ . In addition, the ballot obtained by replacing  $b_{\alpha}$  by j in a given ballot b is denoted  $(b_{-\alpha}, j)$ . Formally,

$$(b_{-\alpha}, j) = \begin{cases} b_{\alpha'} & \text{for all } \alpha' \neq \alpha, \\ j & \text{for } \alpha. \end{cases}$$

We are now able to define the vote shares for each candidate. For any (measurable) set  $C \subseteq \Omega$ , the *vote share* for candidate j, given a ballot  $b \in B$  is given by:

$$v_j^C(b) = \mu(\{\alpha \in C \mid b_\alpha = j\}).$$

We denote the total vote for candidate j by  $v_j(b) = v_j^{\Omega}(b)$ . The vote for a set  $J \subseteq K$  of candidates is given by  $v_J(b) = \sum_{j \in J} v_j(b)$ .

We assume all voters vote sincerely; they cast their ballot for the candidate that they believe is closest to them. We describe the voters' payoff function, given a profile of candidate positions s, as a function  $\pi_{\alpha}: B \to \mathbb{R}$  with  $\pi_{\alpha}(b) = u_{\alpha}(s_{b_{\alpha}})$ . In a complete information setting, this payoff function together with a single-peaked utility function implies sincere voting.

The model distinguishes between voters based on their information. We define these beliefs in the following way. Recall that we let S denote the set of functions from

K into X. These functions represent possible policy positions for the k candidates. Let M(S) denote the set of probability measures on S. Then the belief space of each voter is M(S). Specifically, beliefs of voters are assigned by a ( $\mathcal{F}$  measurable) function  $\lambda:\Omega\to M(S)$ . Thus, voter  $\alpha$ 's belief is given by  $\lambda(\alpha)$ , which is also denoted by  $\lambda_{\alpha}$ , a probability measure on S. The set of all such ( $\mathcal{F}$  measurable) functions from  $\Omega$  to M(S) is denoted by  $\Lambda$ . The beliefs of voters are important because they allow the voters to make decisions based on expected utility calculations when information is incomplete. The expectation operator is defined in the usual way;  $E_{\lambda_{\alpha}}[z] = \int_{s \in S} z(s) d\lambda_{\alpha}$  for (integrable) real-valued functions z with domain S.

In our definition of equilibrium, we include some restrictions about what information is available to the informed and uninformed voters. The informed voters in the model have perfect information about the policy positions of the candidates; their beliefs contain no uncertainty. The uninformed voters, however, do not know the candidates' positions with certainty. Rather, uninformed voters have a specific type of "low-level" knowledge about the policy positions of candidates. Specifically, we suppose that the uninformed voters understand the (ordinal) ordering of the candidates on the issue space. For example, in a three candidate race, while the uninformed voters do not know the exact (spatial) positions of candidates, we suppose they can correctly determine which candidate is on the left of the issue space, which is on the right, and which is the moderate candidate. This limited information could be supplied by cues such as party labels or interest group endorsements or by minimal exposure to the political advertisements or media coverage of a campaign.

In addition, the equilibrium definition implicitly specifies the information that voters observe in the model. All voters, including the uninformed, observe the vote (or poll) shares of the k candidates. In addition, voters know that other voters have symmetric single-peaked utility functions, but they do not know the exact distribution

of voter ideal points. Rather, each voter knows the general features of the distribution such as continuity, monotonicity, etc. as assumed in our results as well as the relative position of the voter's ideal point in the (true) distribution on the issue space. Thus, voter  $\alpha$  knows  $F(y_{\alpha})$  and the general properties of F(x), but not the values of F(x) for  $x \neq y_{\alpha}$ . Finally, we suppose that each voter assumes that all other voters are fully informed and rational. This assumption is justified in equilibrium as all voters act as if they were fully informed.

In the definition of equilibrium, we will employ the following notation. For a given profile of candidate positions s, define  $m_j = (s_j + s_{j+1})/2$  for  $j = 1, \ldots, k-1$  and set  $m_0 = -\infty$  and  $m_k = +\infty$ . Then, under complete information, voters whose ideal points lie in the region between  $m_{j-1}$  and  $m_j$  should vote for candidate j. With this in mind, we define a correspondence  $\varphi: K \times S \to \mathbb{R}$  by  $\varphi(j,s) = \{x \in \mathbb{R} \mid m_{j-1}(s) < x < m_j(s)\}$ . We denote the ideal point region supporting candidate j by  $\varphi_j(s) = \varphi(j,s)$ .

We are now able to formalize the equilibrium concept we use in our analysis. The definition of equilibrium we offer contains two conditions. First, we require that voters cast ballots that maximize their payoffs given their beliefs. Thus each voter implements a best response, conditional on beliefs. Second, we require that the beliefs of each voter must be consistent with the information available to the voter. In other words, voters must have rational expectations; beliefs are self-fulfilling in equilibrium. In addition, our definition includes the assumptions about information discussed above.

**Definition 4.1** Let  $s^* \in S$ . A voter equilibrium, conditional on  $s^*$ , is a pair  $(b^*, \lambda^*) \in B \times \Lambda$  satisfying

1. 
$$\forall \alpha \in \Omega, \ b_{\alpha}^* \ satisfies \max_{j \in K} E_{\lambda_{\alpha}^*}[\pi_{\alpha}(b_{-\alpha}^*, j)] = E_{\lambda_{\alpha}^*}[\pi_{\alpha}(b_{-\alpha}^*, b_{\alpha}^*)],$$

2. 
$$\forall \alpha \in I, \ \lambda_{\alpha}^*(s^*) = 1,$$

3.  $\forall \alpha \in U$ , if  $s^*$  satisfies  $s_i^* \neq s_i^*$  for  $i \neq j$ , then

(a) 
$$\lambda_{\alpha}^*(\{s \in S \mid s_1 < s_2 < \ldots < s_k\}) = 1$$
,

(b) 
$$v_{\{1,\dots,j-1\}}(b^*) < F(y_\alpha^*) < v_{\{1,\dots,j\}}(b^*) \Longrightarrow \lambda_\alpha^*(\{s \in S \mid y_\alpha^* \in \varphi_j(s)\}) = 1.$$

A brief discussion of the three parts of this definition is in order. Condition 4.1.1 states that each voter  $\alpha$  must cast a ballot  $b_{\alpha}^*$  that maximizes the expected value of the payoff function  $\pi_{\alpha}$ , where the expected value is taken over voter  $\alpha$ 's belief  $\lambda_{\alpha}^{*}(s)$ on the candidates' policy positions. For informed voters, condition 4.1.2 requires that these beliefs, which are probability measures, must be point masses on the true candidate positions  $s^*$ . This condition reflects the fact that informed voters know these positions with certainty. For uninformed voters, the definition imposes two conditions in the case that the candidates' positions are all distinct. Condition 4.1.3a reflects the limited knowledge of the uninformed voters about the positions of candidates. Specifically, we require that voter  $\alpha$ 's belief  $\lambda_{\alpha}^*$  assigns zero probability to candidate positions that are not ordered correctly relative to the true positions  $s^*$ . Finally, condition 4.1.3b insures that the beliefs of uninformed voters are in agreement with the observed vote shares. Given that voter  $\alpha$  knows that other voters have symmetric single-peaked utility functions and assumes that all other voters are fully informed and rational, the votes for candidates 1 through j-1 must come from voters with ideal points to the left of  $m_{j-1}(s^*)$ . Likewise, the votes for candidates 1 through j must come from voters with ideal points to the left of  $m_j(s^*)$ . Thus, under these assumptions, if the relative position of the voter's ideal point,  $F(y_{\alpha}^*)$ , lies between  $v_{\{1,\ldots,j-1\}}(b^*)$ , the cumulative vote shares for candidates 1 through j-1, and  $v_{\{1,\ldots,j\}}(b^*)$ , the cumulative vote shares for candidates 1 through j, voter  $\alpha$  can infer

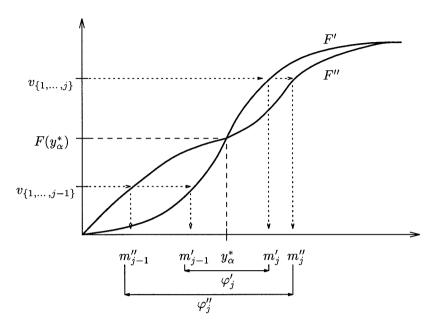


Figure 4.1: Inference of Uninformed Voters

that the candidate positions  $s \in S$  is such that  $m_j(s^*) < y_\alpha^* < m_j(s^*)$ . In this case, voter  $\alpha$ 's belief  $\lambda_\alpha^*$  must assign zero probability to situations in which  $y_\alpha^* \notin \varphi_j(s)$ .

We illustrate this last condition in Figure 4.1 for the case of a election in which  $\Omega$  is infinite and F is invertible. While neither the graph of F(x) or the candidate locations s are available to an uninformed voter  $\alpha$ , the figure illustrates the inference that voter  $\alpha$  is able to make in a voter equilibrium. The horizontal axis represents the issue space X and the vertical axis represents the percentage of voters. Voter  $\alpha$  knows  $y_{\alpha}^*$  and  $F(y_{\alpha}^*)$  and observes the equilibrium vote shares  $v_{\{1,\ldots,j\}}(b^*)$  for all  $j \in K$ . Consider two possible conjectures by voter  $\alpha$  as to the possible distribution of voter ideal points, F' and F'', as indicated in the figure. Note that both these distribution pass through  $F(y_{\alpha}^*)$ , as this information is known to voter  $\alpha$ . Suppose that  $v_{\{1,\ldots,j-1\}}(b^*) < F(y_{\alpha}^*) < v_{\{1,\ldots,j\}}(b^*)$  for some candidate j. Now, a vote share  $v_{\{1,\ldots,j-1\}}$  must imply a candidate midpoint at the point labeled  $m'_{j-1}$  for the conjecture F' and

a midpoint at the point labeled  $m''_{j-1}$  for the conjecture F''. Likewise, conjecture F' implies that  $m'_j$  is the midpoint supporting  $v_{\{1,\ldots,j-1\}}$  and  $m''_j$  for the conjecture F''. These implications then lead to a conjecture  $\varphi'_j$  for the set of voter ideal points voting for candidate j, under the conjectured distribution F', and  $\varphi''_j$  for F'', as illustrated in the figure. These two conjectures make clear the general result — for any possible distribution it must be the case that  $m_{j-1}(s^*)$  lies to the left of  $y^*_\alpha$  and  $m_j(s^*)$  lies to the right. Thus, in equilibrium, voter  $\alpha$ 's beliefs must be that  $y^*_\alpha \in \varphi_j(s)$ .

This inference is the driving force behind the results that we prove. It provides a means by which information can travel from the informed segment of the electorate to the uninformed. Loosely, the inference can be stated as "a candidate whose support comes from moderate voters must also be a moderate."

We end this section with a comment on the requirement that the positions of candidates in the issue space must be distinct. We maintain this restriction in the results presented below. The main purpose of the restriction is to expedite the notation and proofs in the model and similar results could be obtained without this assumption. Moreover, as we showed in the first chapter, a  $\varepsilon$ -separability condition similar to the restriction we impose here is equivalent to several alternative equilibrium concepts for spatial model of elections. Finally, the ability of candidates to run on exactly identical platforms seems empirically unsupportable.

# 4.3 Characterization and Convergence of Voter Equilibria

In this section, we focus on voter equilibria of the model. We state and prove several propositions relating to their characterization and eventual convergence in a dynamic

setting. The first theorem provides an essentially unique characterization of the ballots cast by voters in a voter equilibrium. In the model, informed voters know the policy positions of the candidates and thus are always able to vote for the candidate nearest to their ideal points. Theorem 4.1 proves that, in equilibrium, uninformed voters also vote "correctly" for the candidate nearest them. This result follows from the fact that the uninformed voters' vote decisions, as illustrated in the previous section, are consistent with the resulting vote shares only when the vote decisions are correct almost everywhere. The beliefs of uninformed voters in equilibrium do not equal point masses on the true candidate positions, as is assumed for informed voters. However, these beliefs are "accurate enough" in the sense that they lead uninformed voters to (almost everywhere) vote as if they were fully informed.

Thus, when the hypotheses of the theorem hold, a platform profile  $s^*$  corresponds to a voter equilibrium ballot  $b^*$  that is unique almost everywhere.<sup>3</sup> This justifies references to "the" equilibrium ballot  $b^*$ . We should also note that it is simple to prove existence of a voter equilibria, and thus the theorem is not vacuously true.<sup>4</sup>

**Theorem 4.1** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ , and  $F_I$  and F are continuous and strictly increasing. Let  $s^* \in S$ . Suppose that  $(b^*, \lambda^*) \in B \times \Lambda$  is a voter equilibrium conditional on  $s^*$  and that  $s^*$  satisfies  $s_i^* \neq s_j^*$  for  $i \neq j$ . Then for all  $\alpha \in \Omega$ ,

$$y_{\alpha}^* \in \varphi_j(s^*) \Longrightarrow b_{\alpha}^* = j.$$

<sup>&</sup>lt;sup>3</sup>To say that the voter equilibrium ballot  $b^*$  conditional on  $s^*$  is unique almost everywhere means that if we denote the set of all voter equilibrium ballots  $b^*$  conditional on  $s^*$  by  $B^*$ , then  $\mu\left(\left\{\alpha\in\Omega\mid\exists b_\alpha\in K\text{ such that }b_\alpha=b_\alpha^*\forall b^*\in B^*\right\}\right)=1$ . This follows directly from the theorem and the fact that  $\bigcup_{j=1}^k\varphi(j,s)=X\setminus\{m_0,m_1,\ldots,m_k\}$  and F is assumed to be nonatomic.

<sup>&</sup>lt;sup>4</sup>For  $s^*$  satisfying  $s_i^* \neq s_j^*$  for  $i \neq j$ , let b be as specified in the theorem and let  $\lambda$  satisfy  $\forall \alpha \in \Omega$ ,  $\lambda_{\alpha}(s^*) = 1$ . Then  $(b, \lambda)$  is a voter equilibrium conditional on  $s^*$ .

Proof: Definition 4.1(1) simplifies to the condition that

$$\forall \alpha \in \Omega, b_{\alpha}^* \in \arg \max_{b_{\alpha} \in K} E_{\lambda_{\alpha}^*}[u_{\alpha}(s_{b_{\alpha}})].$$

As  $u_{\alpha}$  is symmetric and single-peaked about  $y_{\alpha}^*$ ,  $u_{\alpha}(s_{b_{\alpha}})$  is maximized by  $b_{\alpha}=j$  if  $y_{\alpha}^* \in \varphi_j(s)$ .

For  $\alpha \in I$ , Definition 4.1.2 implies that

$$E_{\lambda_{\alpha}^*}[u_{\alpha}(s_{b_{\alpha}})] = u_{\alpha}(s_{b_{\alpha}}^*).$$

Thus, for  $\alpha \in I$ , the result of the theorem holds.

For  $\alpha \in U$ , suppose that  $v_{\{1,\ldots,j-1\}}(b^*) < F(y_{\alpha}^*) < v_{\{1,\ldots,j\}}(b^*)$  for some  $j \in K$  (this holds except for a  $\mu$ -measure zero set). Then Definitions 4.1.3a and 4.1.3b imply that

$$E_{\lambda_{\alpha}^*}[u_{\alpha}(s_{b_{\alpha}})] = \int_{\{s \in S \mid y_{\alpha}^* \in \varphi_i(s)\}} u_{\alpha}(s_{b_{\alpha}}) d\lambda_{\alpha}^*.$$

But as the integrand is maximized by  $b_{\alpha} = j$  everywhere in the region of integration, the integral must also be maximized at  $b_{\alpha} = j$ . We conclude that for  $\alpha \in U$ ,

$$v_{\{1,\dots,j-1\}}(b^*) < F(y_\alpha^*) < v_{\{1,\dots,j\}}(b^*) \Longrightarrow b_\alpha^* = j.$$

As F is continuous and strictly increasing,  $F^{-1}$  exists and is strictly increasing. In order to complete the proof, define  $x_j = F^{-1}(v_{\{1,\dots,j\}}(b^*))$  for  $j \in K$  and define  $x_0 = F^{-1}(0)$ . By monotonicity of F, it follows that for all  $j \in K$ ,

$$x_{j-1} < y_{\alpha}^* < x_j \implies F(x_{j-1}) < F(y_{\alpha}^*) < F(x_j)$$

$$\implies v_{\{1,\dots,j-1\}}(b^*) < F(y_{\alpha}^*) < v_{\{1,\dots,j\}}(b^*)$$

$$\implies b_{\alpha}^* = j.$$

We can thus give a lower bound on the vote for candidate j. For all  $j \in K$ ,

$$v_j(b^*) \ge F_I(m_j) - F_I(m_{j-1}) + F_U(x_j) - F_U(x_{j-1}).$$

If j = 1 this simplifies to  $v_1(b^*) \ge F_I(m_1) + F_U(x_1)$ .

As  $v_{\{1,...,j\}}(b^*) = \sum_{i=1}^{j} v_i(b^*)$  for  $j \in K$ , it follows that

$$v_{\{1,\dots,j\}}(b^*) \geq \sum_{i=1}^{j} F_I(m_i) - F_I(m_{i-1}) + F_U(x_i) - F_U(x_{i-1})$$

$$\geq (F_I(m_j) + F_U(x_j)) - (F_I(m_{j-1}) + F_U(x_{j-1})) +$$

$$(F_I(m_{j-1}) + F_U(x_{j-1})) - (F_I(m_{j-2}) + F_U(x_{j-2})) +$$

$$\cdots + (F_I(m_1) + F_U(x_1)) - (F_I(m_1) + F_U(x_1))$$

$$\geq F_I(m_j) + F_U(x_j)$$

$$\geq F_I(m_j) - F_I(x_j) + (F_I(x_j) + F_U(x_j))$$

$$\geq F_I(m_j) - F_I(x_j) + F_I(x_j)$$

Thus  $F_I(m_j) \leq F_I(x_j)$  holds for all  $j \in K$ .

Note also that

$$v_k(b^*) \geq F_I(m_k) - F_I(m_{k-1}) + F_U(x_k) - F_U(x_{k-1})$$

$$\geq 1 - F_I(m_{k-1}) - F_U(x_{k-1})$$

$$\geq 1 - F_I(m_{k-1}) + F_I(x_{k-1}) - (F_I(x_{k-1}) - F_U(x_{k-1}))$$

$$\geq 1 - F_I(m_{k-1}) + F_I(x_{k-1}) - F(x_{k-1})$$

$$\geq 1 - F_I(m_{k-1}) + F_I(x_{k-1}) - F(x_{k-1})$$

$$\geq 1 - F_I(m_{k-1}) + F_I(x_{k-1}) - F(x_{k-1})$$

$$v_k(b^*) + v_{\{1,\dots,k-1\}}(b^*) \geq 1 - F_I(m_{k-1}) + F_I(x_{k-1})$$

$$0 \geq F_I(x_{k-1}) - F_I(m_{k-1}).$$

This implies that  $F_I(x_{k-1}) \leq F_I(m_{k-1})$ . A similar argument shows that  $F_I(x_j) \leq F_I(m_j)$  for all  $j \in K$ .

From  $F_I(x_j) \leq F_I(m_j) \leq F_I(x_j)$ , it follows that  $F_I(x_j) = F_I(m_j)$  for all  $j \in K$ . As  $F_I$  is invertible,  $x_j = m_j$  for all  $j \in K$ . Thus, for  $\alpha \in \Omega$ ,

$$y_{\alpha}^* \in \varphi_j(s^*) \Longrightarrow v_{\{1,\dots,j-1\}}(b^*) < F(y_{\alpha}^*) < v_{\{1,\dots,j\}}(b^*) \Longrightarrow b_{\alpha}^* = j.$$

We thus conclude that, in equilibrium, both informed and uninformed voters cast ballots as if they had correct information about the positions of candidates. While the beliefs of the uninformed voters may remain imprecise, they still are accurate enough to allow the uninformed voters to vote correctly.

Theorem 4.1 proves that "correct" voting by all voters is an equilibrium state; all voters are maximizing payoffs subject to beliefs and these beliefs are consistent with the observed data. But the theorem is silent about how such an state might be reached. Certainly the merits of this equilibrium are irrelevant unless it is achievable. In order to address this concern, we next construct a dynamic process, corresponding

to a sequence of successive polls, which converges to this voter equilibrium for any initial configuration of beliefs.

We formally define this dynamic process as follows. We begin with an arbitrary initial ballot  $b^0$  and belief  $\lambda^0$ . Now suppose a series of consecutive polls is conducted, with each poll's results announced publicly before the next one is taken. As each poll is a survey of the entire population of voters rather than a random selection from the population, we also use the term "straw vote." As we specify in definition 4.2, all voters respond to polls rationally in the sense that they adjust their responses to be consistent with the information reflected in the most recently announced poll. Specifically, we suppose that for the poll in period t, all voters cast a straw vote that is optimal according to their beliefs  $\lambda^{t-1}$  and form beliefs consistent with the latest results,  $v_{\{1,\dots,j\}}(b^{t-1})$  for all  $j \in K$ .

**Definition 4.2** Let  $s^* \in S$  and fix a pair  $(b^0, \lambda^0) \in B \times \Lambda$ . For t = 1, 2, ..., we recursively define a sequence  $\{(b^t, \lambda^t)\}$ , where  $(b^t, \lambda^t) \in B \times \Lambda$  for all t = 1, 2, ..., by

1. 
$$\forall \alpha \in \Omega, \ b_{\alpha}^{t} \ satisfies \max_{j \in K} E_{\lambda_{\alpha}^{t-1}}[\pi_{\alpha}(b,j)] = E_{\lambda_{\alpha}^{t-1}}[\pi_{\alpha}(b,b_{\alpha}^{t})],$$

- 2.  $\forall \alpha \in I, \ \lambda_{\alpha}^{t}(s^{*}) = 1,$
- 3.  $\forall \alpha \in U$ , if  $s^*$  satisfies  $s_i^* \neq s_j^*$  for  $i \neq j$ , then

(a) 
$$\lambda_{\alpha}^{t}(\{s \in S \mid s_{1} < s_{2} < \ldots < s_{k}\}) = 1,$$

$$(b) \ v_{\{1,\dots,j-1\}}(b^{t-1}) < F(y_{\alpha}^*) < v_{\{1,\dots,j\}}(b^{t-1}) \Longrightarrow \lambda_{\alpha}^t(\{s \in S \mid y_{\alpha}^* \in \varphi_j(s)\}) = 1.$$

It follows from the proof of Theorem 4.1 that if the assumptions of the theorem hold, the ballots  $b^t$  defined by Definition 4.2 satisfy, for all  $j \in K$  and t = 1, 2, ...,

$$\alpha \in I \implies m_{j-1}(s^*) < y_{\alpha}^* < m_j(s^*) \Longrightarrow b_{\alpha}^t = j$$
 (4.1)

$$\alpha \in U \implies x_{j-1}^{t-1} < y_{\alpha}^* < x_j^{t-1} \Longrightarrow b_{\alpha}^t = j, \tag{4.2}$$

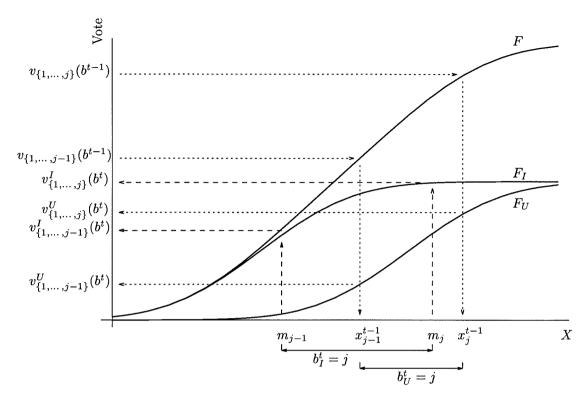


Figure 4.2: Vote Response In Successive Polls

where  $x_j^t = F^{-1}(v_{\{1,\ldots,j\}}(b^t))$  for  $j \in K$  and  $x_0^t = F^{-1}(0)$ , for  $t = 1, 2, \ldots$  These conditions dictate what ballots voters cast in succeeding polls as a function of most recent poll's information.

Figure 4.2 illustrates voters' responses in successive polls. Again the horizontal axis represents the issue space X and the vertical axis represents vote shares. After the poll at time t-1, the results, vote shares for all candidates, are publicly announced. With this information, the uninformed voters are able to calculate  $x_j^{t-1} = F^{-1}(v_{\{1,\ldots,j\}}(b^{t-1}))$  for  $j \in K$ , as indicated in the figure. Then, for the poll at time t, uninformed voters with ideal points between  $x_{j-1}^{t-1}$  and  $x_j^{t-1}$  vote for candidate j. Of course informed voters know the true candidate positions in all periods and cast correct votes in all polls. The resulting vote shares for the time t poll are identified in the figure for both groups of voters.

The following theorem shows that even from an arbitrary initial condition, the above process converges to a voter equilibrium. In other words, the eventual outcome of the dynamic polling process is that all voters vote as if they were fully informed about the positions of the candidates.

**Theorem 4.2** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ , and  $F_I$ ,  $F_U$  and F are continuous and strictly increasing. If  $s^* \in S$  satisfies  $s_i^* \neq s_j^*$  for  $i \neq j$ , then the process defined by Definition 4.2 converges to the unique (up to a set of measure 0) voter equilibrium ballot  $b^*$  corresponding to  $s^*$ , in the sense that  $v_j(b^t) \to v_j(b^*)$  and  $x_j^t \to m_j(s^*)$  for all  $j \in K$  as  $t \to \infty$ .

*Proof*: As F is continuous and strictly increasing, if we show that  $x_j^t \to m_j(s^*)$  for all  $j \in K$  as  $t \to \infty$ , then for all  $\alpha \in \Omega$  and  $j \in K$ ,

$$y_{\alpha}^* \in \varphi_j(s^*) \Longrightarrow b_{\alpha}^* = j$$

and thus  $v_j(b^t) \to v_j(b^*)$  for all  $j \in K$  as  $t \to \infty$ .

To begin the proof, focus on the vote for candidate 1. From equations 4.1, we have

$$v_1(b^t) = F_I(m_1) + F_U(x_1^{t-1})$$

$$= F_I(m_1) - F_I(x_1^{t-1}) + (F_I(x_1^{t-1}) + F_U(x_1^{t-1}))$$

$$= F_I(m_1) - F_I(x_1^{t-1}) + F(x_1^{t-1}).$$

Consider first the case  $x_1^t < m_1(s^*)$ . Then, as  $F_I$  is strictly increasing,  $F_I(m_1) - F_I(x_1^{t-1}) > 0$ . As F is continuous and strictly increasing,  $F^{-1}$  is continuous and

strictly increasing. Thus,

$$x_1^t = F^{-1}(v_1(b^t)) = F^{-1}\left(F(x_1^{t-1}) + F_I(m_1) - F_I(x_1^{t-1})\right) > F^{-1}\left(F(x_1^{t-1})\right) = x_1^{t-1}.$$

We can also write

$$v_1(b^t) = F_I(m_1) + F_U(x_1^{t-1}) = F(m_1) - [F_U(m_1) - F_U(x_1^{t-1})].$$

As  $F_U$  is strictly increasing,  $F_U(m_1) - F_U(x_1^{t-1}) > 0$ . Thus,

$$x_1^t = F^{-1}(v_1(b^t)) = F^{-1}\left(F(m_1) - \left[F_U(m_1) - F_U(x_1^{t-1})\right]\right)$$

$$< F^{-1}\left(F(m_1(s^*))\right) = m_1(s^*).$$

So if  $x_1^0 < m_1(s^*)$ , the sequence  $\{x_1^t\}$  has the property that  $x_1^{t-1} < x_1^t < m_1(s^*)$ . So  $\{x_1^t\}$  is a bounded monotonic sequence and thus must converge to its least upper bound. We claim that this least upper bound is  $m_1(s^*)$ . To show this, suppose that the least upper bound of the sequence  $\{x_1^t\}$  is  $\bar{x} < m_1(s^*)$ . Then as  $F_I$  is strictly increasing,  $F_I(m_1(s^*)) - F_I(\bar{x}) > 0$ . As F is continuous and strictly increasing, there exists an integer T such that t > T implies  $F(\bar{x}) - F(x_1^t) < F_I(m_1(s^*)) - F_I(\bar{x})$ . It follows from  $x_1^t < \bar{x}$  that  $F_I(x_1^t) < F_I(\bar{x})$  and thus

$$\bar{x} = F^{-1}(F(\bar{x})) < F^{-1}(F(x_1^t) + F_I(m_1) - F_I(\bar{x}))$$

$$< F^{-1}(F(x_1^t) + F_I(m_1) - F_I(x_1^t)) = x_1^{t+1}.$$

This contradicts the assumption that  $\bar{x}$  is an upper bound, so the sequence  $\{x_1^t\}$  must converge to  $m_1(s^*)$ .

A similar argument applied to the case  $x_1^t > m_1(s^*)$  proves that if  $x_1^0 > m_1(s^*)$ , the sequence  $\{x_1^t\}$  is strictly decreasing and also converges to  $m_1(s^*)$ . In addition, this method of proof applies to candidate k and yields the result that  $x_k^t \to m_k(s^*)$  monotonically as  $t \to \infty$ .

To prove the result for  $j=2,\ldots,k-1$ , we have from equations 4.1 that

$$\begin{split} v_{\{1,\ldots,j\}}(b^t) &= \sum_{i=1}^j v_i(b^t) \\ &= \sum_{i=1}^j F_I(m_i) - F_I(m_{i-1}) + F_U(x_i^{t-1}) - F_U(x_{i-1}^{t-1}) \\ &= (F_I(m_j) + F_U(x_j^{t-1})) - (F_I(m_{j-1}) + F_U(x_{j-1}^{t-1})) + \\ (F_I(m_{j-1}) + F_U(x_{j-1}^{t-1})) - (F_I(m_{j-2}) + F_U(x_{j-2}^{t-1})) + \\ &\cdots + (F_I(m_1) + F_U(x_1^{t-1})) - (F_I(m_1) + F_U(x_1^{t-1})) \\ &= F_I(m_j) + F_U(x_j^{t-1}) \\ &= F_I(m_j) - F_I(x_j^{t-1}) + (F_I(x_j^{t-1}) + F_U(x_j^{t-1})) \\ &= F(x_j^{t-1}) + F_I(m_j) - F_I(x_j^{t-1}). \end{split}$$

If we suppose that  $x_j^t < m_j(s^*)$ , then  $F_I(m_j) - F_I(x_j^{t-1}) > 0$  and thus

$$x_j^t = F^{-1}(v_{\{1,\ldots,j\}}(b^t)) > F^{-1}\left(F(x_j^{t-1})\right) = x_j^{t-1}.$$

Rewriting  $v_{\{1,\dots,j\}}(b^t)$  as  $F(m_j)-\left(F_U(m_j)-F_U(x_j^{t-1})\right)$  and noting that  $F_U(m_j)-F_U(x_j^{t-1})>0$  yields

$$x_i^t = F^{-1}(v_{\{1,\dots,i\}}(b^t)) < F^{-1}(F(m_i(s^*))) = m_i(s^*).$$

Repeating the argument on least upper bounds proves that the sequence  $\{x_j^t\}$  is

monotonically increasing and converges to  $m_j(s^*)$ . When  $x_j^t > m_j(s^*)$ , the sequence  $\{x_j^t\}$  is strictly decreasing and also converges to  $m_j(s^*)$ .

We therefore conclude that  $v_j(b^t) \to v_j(b^*)$  and  $x_j^t \to m_j(s^*)$  for all  $j \in K$  as  $t \to \infty$ .

## 4.4 Full Equilibria

Up to this point, we have assumed that the candidates' positions in the model are fixed exogenously and concentrated on the information and decisions of the voters. Under these assumptions, we have shown that in a voter equilibrium all voters act as if they had complete information. We now expand the model to include the candidates as strategic actors with incomplete information about the characteristics of the voters. In this section we show that the candidates, as well as voters, act in equilibrium as if they were fully informed.

We maintain all of the relevant notation from section 4.2 and also define some new concepts below. Recall that S denotes the set of functions from K into X and a typical element  $s(j) = s_j$  represents a policy position for candidate j. For the strategy profile obtained by replacing  $s_j$  by  $s'_j$  in s, we use the notation  $(s_{-j}, s'_j)$ . In addition, to emphasize the dependence of voters' ballots on the policy positions of the candidates in equilibrium, we sometimes write b(s) for the ballot b.

We suppose that candidates choose their strategies in an attempt to maximize the vote share they receive on election day. Therefore, we define candidate j's payoff function  $\pi_j: S \times B \to \mathbb{R}$  as a function of a profile of policy positions s and a ballot b(s) which yield vote shares. Specifically, we define

$$\pi_j(s,b) = v_j(b(s)).$$

We next define the information and beliefs of the candidates. As with voters, candidates' beliefs will be specified as probability measures over the possible distributions of voter characteristics. We assume that each candidate knows the policy positions of all k candidates. We denote the set of possible distributions of ideal points for voters by  $Y = \{y \mid y : \Omega \to X \text{ is } \mathcal{F} \text{ measurable}\}$ . For any such function y, the corresponding distribution function is defined as  $F_C(x \mid y) = \mu(\{\alpha \in C \mid y_\alpha \leq x\})$  for any (measurable) set  $C \subseteq \Omega$  and  $x \in \mathbb{R}$ . The belief space of each candidate is M(Y), the set of probability measures on some  $\sigma$  algebra of Y. Specifically, candidates' beliefs are assigned by a function  $\gamma: K \to M(Y)$ . Thus, candidate j's belief about voter ideal points is given by  $\gamma(j)$ , which is also denoted by  $\gamma_j$ , a probability measure on some  $\sigma$  algebra of Y. The set of all such belief functions from K to M(Y) is denoted by  $\Gamma$ .

As with voters, a candidate's beliefs allow the candidate to make expected utility calculations in the face of incomplete information about the voters in the model. We suppose that candidates assume that voters are fully informed and rational;<sup>6</sup> they vote (sincerely) for the candidate closest to their ideal point. For a possible distribution of voter ideal points y, candidates can construct a *predicted ballot* for a profile of candidate policy positions s. For  $s \in S$  and  $y \in Y$ , we define the predicted ballot  $\hat{b}(s,y) \in B$  by

$$\hat{b}_{\alpha}(s,y) = \begin{cases} j & \text{if } u(s_{j}, y_{\alpha}) > u(s_{j'}, y_{\alpha}) \text{ for all } j' \neq j, \\ 1 & \text{otherwise,} \end{cases}$$

<sup>&</sup>lt;sup>5</sup>We assume that for all  $x, t \in \mathbb{R}$  and  $C \in \Omega$ , all sets of the form  $\{y \in Y \mid F_C(x \mid y) \geq t\}$  are Y measurable.

<sup>&</sup>lt;sup>6</sup>As we have noted, this assumption about voters is not correct in general. However, *in equilibrium* this assumption is justified, as all voters act as if they were fully informed.

for each  $\alpha \in \Omega$ . As was the case in an earlier section,  $\hat{b}(s,y)$  will be unique almost everywhere, while the predicted vote totals for  $\hat{b}(s,y)$  will be unique.

As in section 4.3, an equilibrium of this model requires that players maximize payoffs conditional on beliefs which are also consistent with available information. But now the set of players is expanded to include both voters and candidates. Thus, a full equilibrium consists of a voter equilibrium conditional on policy positions that, along with beliefs, form a candidate equilibrium.

**Definition 4.3** A full equilibrium is an element  $(s^*, \gamma^*, b^*, \lambda^*) \in S \times \Gamma \times B \times \Lambda$  such that  $(b^*, \lambda^*)$  is a voter equilibrium, conditional on  $s^*$ , and for all  $j \in K$ , the pair  $(s^*, \gamma^*)$  satisfies

$$1. \ s_j^* \in \arg\max_{s_j' \in X} E_{\gamma_j^*}[\pi_j((s_{-j}^*, s_j'), \hat{b}((s_{-j}^*, s_j'), y))],$$

2. for all 
$$i \in K$$
,  $\gamma_j^* (\{y \in Y \mid F(m_i(s^*) \mid y) = v_{\{1,\dots,i\}}(b^*)\}) = 1$ .

Before proceeding, we briefly review the components of this definition. Condition 4.3.1 states that each candidate j must adopt a policy position  $s_j^*$  that maximizes the expected value of the payoff function  $\pi_j$ , where the expected value is taken over candidate j's belief  $\gamma_j^*(y)$  on the distribution of voter ideal points. In addition, condition 4.3.2 requires that these beliefs be consistent with the observed vote shares. Candidate j knows the policy positions advocated by the other candidates and can thus determine the policy midpoints  $m_i(s^*)$  for all  $i \in K$ . Candidate j also assumes that the voters are also fully informed about  $s^*$  and all vote correctly for the nearest candidate, which is true almost everywhere in a voter equilibrium. Thus, for each  $i \in K$ , the combined vote shares for candidates  $1, \ldots, i$  must equal the share of voters to the left of  $m_i(s^*)$ . Therefore candidate j's belief  $\gamma_j^*$  must assign zero probability to distributions of voter ideal points y in which  $F(m_i(s^*) \mid y) \neq v_{\{1,\ldots,i\}}(b^*)$  for any

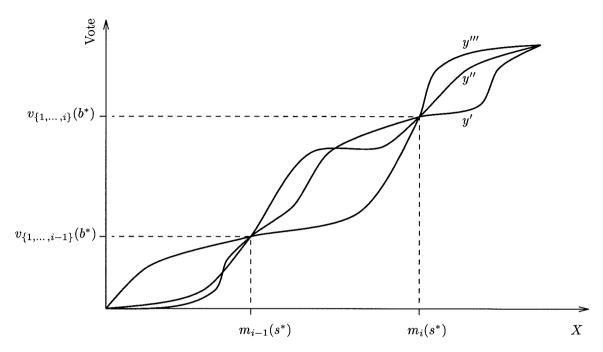


Figure 4.3: Candidate Beliefs In A Full Equilibrium

 $i \in K$ . In other words, the distribution function  $F(x \mid y)$  for each each admissible distribution y must have value  $v_{\{1,\ldots,i\}}(b^*)$  at  $m_i(s^*)$ , as illustrated in Figure 4.3 for three possible distributions y', y'', and y'''.

In order to illustrate the characteristics of full equilibria in a simple setting, we consider a three-candidate election. The first theorem we prove shows that no full equilibrium exists in this case; a result analogous to the complete information setting.

**Theorem 4.3** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ ,  $F_I(x \mid y)$ ,  $F_U(x \mid y)$  and  $F(x \mid y)$  are continuous and strictly increasing for all  $y \in Y$ , and k = 3. Then no full equilibrium  $(s^*, \gamma^*, b^*, \lambda^*)$  exists in which  $s^*$  satisfies  $s_i^* \neq s_j^*$  for  $i \neq j$ .

*Proof*: From the definition of a predicted ballot  $\hat{b}$  and of the candidates payoff function  $\pi_j(s,b)$ , it follows that for any profile s and any possible distribution of voter ideal

points y,

$$\pi_j((s_{-j}^*, s_j'), \hat{b}((s_{-j}^*, s_j'), y)) = v_j(\hat{b}((s_{-j}^*, s_j'), y))$$

$$= F(m_j((s_{-j}^*, s_j')) \mid y) - F(m_{j-1}((s_{-j}^*, s_j')) \mid y).$$

Consider candidate 1. We have

$$\pi_1((s_{-1}^*, s_1'), \hat{b}((s_{-1}^*, s_1'), y)) = F(m_1((s_{-1}^*, s_1')) \mid y).$$

Suppose that y satisfies  $F(m_1(s^*) \mid y) = v_1(b^*)$ . Then as  $F(x \mid y)$  is strictly increasing, for some  $\varepsilon > 0$  such that  $s_1' = s_1^* + \varepsilon < s_2^*$ ,

$$F(m_1((s_{-1}^*, s_1')) \mid y) > F(m_1(s^*) \mid y).$$

As this holds for all  $\{y \in Y \mid F(m_1(s^*) \mid y) = v_1(b^*)\}$ , definition 4.3.2 implies that

$$E_{\gamma_i^*}[\pi_1((s_{-1}^*, s_1'), \hat{b}((s_{-1}^*, s_1'), y))] > E_{\gamma_i^*}[\pi_1(s^*, \hat{b}(s^*, y))].$$

This contradicts the maximization condition in definition 4.3.1, so no such equilibrium exists.

It is also clear that no full equilibrium exists in which any of three candidates adopt identical positions.

It may appear that this result simply reflects the lack of best response for an outside candidate. The next result shows that this is not the case. In order to guarantee the existence of a best response, we assume a  $\varepsilon$ -separability condition similar to that presented in chapter 2. **Definition 4.4** For a given  $\varepsilon > 0$ ,  $s \in S$  satisfies  $\varepsilon$ -separability if  $|s_i - s_j| \ge \varepsilon$  for  $i \ne j$ .

We next define a  $\varepsilon$ -separable equilibrium as the limit of a sequence of full equilibria satisfying  $\varepsilon$ -separability.

**Definition 4.5** An element  $(s^*, \gamma^*, b^*, \lambda^*) \in S \times \Gamma \times B \times \Lambda$  is an  $\varepsilon$ -separable equilibrium if there exists a sequence  $\{(s^{\varepsilon}, \gamma^{\varepsilon}, b^{\varepsilon}, \lambda^{\varepsilon*})\}$  such that  $s^{\varepsilon} \to s^*$  as  $\varepsilon \to 0$ ,  $s^{\varepsilon}$  satisfies  $\varepsilon$ -separability,  $\gamma^{\varepsilon} = \gamma^*$  for all  $\varepsilon > 0$ ,  $(b^{\varepsilon}, \lambda^{\varepsilon})$  is a voter equilibrium, conditional on  $s^{\varepsilon}$ , and for all  $j \in K$ , the pair  $(s^{\varepsilon}, \gamma^{\varepsilon})$  satisfies

1. 
$$s_j^{\varepsilon} \in \arg\max_{s_i' \in X_{-i}^{\varepsilon}} E_{\gamma_i^{\varepsilon}} [\pi_j((s_{-j}^{\varepsilon}, s_j'), \hat{b}((s_{-j}^{\varepsilon}, s_j'), y))],$$

2. for all 
$$i \in K$$
,  $\gamma_j^{\varepsilon} \left( \left\{ y \in Y \mid F(m_i(s^{\varepsilon}) \mid y) = v_{\{1,\dots,i\}}(b^{\varepsilon}) \right\} \right) = 1$ .

where 
$$X_{-j}^{\varepsilon} = X \setminus \bigcup_{i \neq j} (s_i^{\varepsilon} - \varepsilon, s_i^{\varepsilon} + \varepsilon).$$

We state the following theorem without proof.

**Theorem 4.4** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ ,  $F_I(x \mid y)$ ,  $F_U(x \mid y)$  and  $F(x \mid y)$  are continuous and strictly increasing for all  $y \in Y$ , and k = 3. Then no  $\varepsilon$ -separable equilibrium exists.

However, as is the case in a complete information setting, if we recast this three candidate competition as spatial competition with entry, then a full equilibrium exists. Palfrey (1984) proves this result for complete information and a similar argument applies here. In a model with entry, two "dominant" candidates choose positions in the issue space and then an entrant chooses a position, knowing the positions of the preceding candidates.

Because of the added component of the timing of candidates' choices, we must modify some of the elements of our model. First, we reinterpret the assumption that  $s_1 \leq s_2 \leq s_3$  by referring to the two established candidates as A and B and the entrant as candidate C and then relabeling them after their choices such that  $s_1 \leq s_2 \leq s_3$ . Using this notation, we suppose that for  $s_A, s_B \in X$ , the entrant's position,  $s_C(s_A, s_B) \in X$  is a (set) function of the dominant candidates positions. In addition, the later move by the entrant changes the information structure of the game played by candidates. We now assume that for all  $j \in \{A, B, C\}$ , and for each  $i \in \{A, B\}$ ,

$$\gamma_i^* (\{ y \in Y \mid F(s_i \mid y) = F(s_i \mid y_\alpha^*) \}) = 1,$$

$$\gamma_j^* \left( \left\{ y \in Y \mid F((s_i + \frac{1}{2})/2 \mid y) = F((s_i + \frac{1}{2})/2 \mid y_\alpha^*) \right\} \right) = 1.$$

In other words, all candidates are aware of the vote shares available to the entrant on the left, right, or center of the distribution. As it turns out, the values specified correspond to the possible optimal entry points of the entrant. In addition, to simplify the presentation we suppose that X = [0,1] and F(x) is continuous, strictly increasing, symmetric, and convex on  $[0,\frac{1}{2}]$  and concave on  $[\frac{1}{2},1]$ . These assumptions correspond to Assumptions 2.1,2.2, and 2.3. We also suppose that the candidates beliefs admit only distributions  $y \in Y$  whose distribution functions satisfy these assumptions.

We are then able to prove

**Theorem 4.5** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ ,  $F_I(x \mid y)$ ,  $F_U(x \mid y)$  and  $F(x \mid y)$  are continuous and strictly increasing for all  $y \in Y$ , and k = 3. Then there exists a  $\varepsilon$ -separable equilibrium with entry  $((s_A^*, s_B^*, s_C^*(s_A^*, s_B^*)), \gamma^*, b^*, \lambda^*)$  that satisfies  $F(s_A) = 1 - 2F(1/4 + s_A/2)$  and  $s_B = 1 - s_A$ .

*Proof*: The proof is similar to much of the material in chapter 2. Fix  $\varepsilon > 0$  and

suppose that  $s_A \leq s_B$ . First, regardless of the entrant's beliefs on  $y \in Y$ , as  $F(x \mid y)$  is strictly increasing, the best response of the entrant to the left of A is  $s_A - \varepsilon$  and the best response of the entrant to the right of B is  $s_B + \varepsilon$ . Since  $F(x \mid y)$  is symmetric and convex on  $[0, \frac{1}{2}]$  and concave on  $[\frac{1}{2}, 1]$ , the optimal entry position between A and B is 1/2 if  $s_B = 1 - s_A$ .

From the results in chapter 2 we know that  $F(s_A^*) = 1 - 2F(1/4 + s_A^*/2)$  and  $s_B^* = 1 - s_A^*$  implies

$$F(s_A^*) = F((s_B^* + 1/2)/2) - F((s_A^* + 1/2)/2).$$

The term on the left side of this equation is the vote share received by C for (optimally) entering on the left, in the limit. The term on the right side of this equation is the vote share received by C for (optimally) entering in the middle, in the limit. Now consider candidate A. If he repositions himself to the left of  $s_A^*$ , candidate C will enter in the middle to maximize votes and by the proof of theorem 4.3 candidate C will receive a smaller vote share. If candidate A chooses a position to the right of  $s_A^*$ , the entrant will enter to the left. Regardless of A's belief on y, A knows that  $v_A(s^*) > F(s_A^* \mid y) = F(s_A^*)$ . But A's beliefs about  $F(x \mid y)$  also guarantee that for  $s_A^* < s_A < s_B^*$ ,  $v_A((s_A, s_B^*) < F((s_B^* + 1/2)/2) - F((s_A^* + 1/2)/2)$ . Thus  $v_A((s_A, s_B^*) < F(s_A^*) < v_A(s^*)$ . Finally it is clear that for  $s_A > s_B^*$ ,  $v_A((s_A, s_B^*) < 1 - F(s_B^*) = v_A(s^*)$ .

Therefore candidate A has no incentive to change his position. A similar argument can be made for candidate B and the proof is complete.

Thus, even when candidates have limited information, the threat of entry by a third-party candidate forces the dominant candidates to adopt non-convergent equilibrium policy positions.

## 4.5 Strategic Voting and Bandwagons

In this section, we suppose that the informed voters in the models vote strategically, as well as being fully informed. We show that this results in an equilibrium in which all voters, including uninformed sincere voters, act as if they were voting strategically based on full information. We conclude that the "wasted vote" phenomenon may thus be the result of information shortcuts used by voters with limited information.

If the informed voters anticipate the results of the dynamic polling process presented in section 4.3, a fraction of this group will realize that their votes will be "wasted" on losing candidates. These voters will then have an incentive to vote for a less preferred candidate who might have a chance of winning. When this occurs, the inferences that the uninformed voters are making are changed, and the eventual result is that *all* voters, informed and uninformed, act as if they were voting strategically. So the process we describe results in uninformed sincere voters voting as if they were fully informed strategic voters.

Suppose that for fear of a wasted vote, informed voters cast ballots for their most preferred viable candidate, where viable is defined as a candidate who receives the largest or second largest share of votes or a candidate tied for first in the case of a tie. The vote shares that informed voters use in this determination are "correct" in the following sense. We suppose that informed voters know the true distribution of voter preferences, F(x), as well as the true candidate positions. As in the earlier sections, we suppose that the uninformed voters believe that all voters are fully informed and vote sincerely. Given this belief and their knowledge of F(x), informed voters can form a shared prediction of the vote shares to be received by a fully informed, sincere electorate. Thus, each informed voter reacts optimally to a sincere electorate, neglecting the other informed voters who are making similar evaluations.

Of course, an objection that immediately arises from those schooled in the socalled "paradox of voting" is that if  $\Omega$  is infinite then no single voter can change the outcome of the election. Therefore, either all or none of the voters are "wasting" their votes by voting sincerely, depending on how you look at it. But we can suppose that the model just described is the limiting case of a finite sequence of models, each of which contains a finite set of voters that must make strategic voting calculations. In fact, just such an argument is the basis for a recent paper on voting equilibria (Myerson and Weber 1993), as well as the results in chapter 3 on strategic voting in large electorates. Myerson and Weber impose an "ordering condition" that requires voters to judge that "[trailing candidates are] much less likely to be serious contenders for first place than candidates with higher predicted [vote shares]." (p. 104) They show that if the ordering condition holds as the size of the electorate grows without bound, then voters will only vote for a candidate who is tied for first, is the unique likely winner, or has the second-largest vote share in the case of a unique likely winner.<sup>7</sup>

This result justifies the following specification for the payoff function of informed voters. First, we define the set of winning candidates for a ballot  $b \in B$  and positions s by

$$w(s,b) = \{j \in K \mid v_j(b(s)) = \max_{i \in K} v_i(b(s))\}.$$

It is clear that  $w(s,b) \subseteq K$  and that |w(s,b)| > 1 implies there is a (|w(s,b)|-way) tie

<sup>&</sup>lt;sup>7</sup>This result is Theorem 2 in Myerson and Weber (1993).

for first. Then, for all  $\alpha \in I$  we define

$$\pi_{\alpha}(b) = \begin{cases} u_{\alpha}(s_{b_{\alpha}}) & \text{if } b_{\alpha} \in w(s, b) \text{ or} \\ |w(p)| = 1 \text{ and } v_{b_{\alpha}} \ge v_{j} \text{ for } j \in K \setminus w(p), \\ -M & \text{otherwise,} \end{cases}$$

for a sufficiently large integer M. Finally, we define expected sincere vote shares for the candidates by  $\hat{v}_j(s^*) = F(m_j(s^*)) - F(m_{j-1}(s^*))$  for all  $j \in K$ . The theorem we present next deals with the case of a single winner and runner-up, for ease of exposition. Similar results can also be proved for the other cases mentioned above.

**Theorem 4.6** Assume  $\Omega$  is infinite,  $\mu(I) > 0$ , and  $F_I$  and F are continuous and strictly increasing. Let  $s^* \in S$ . Suppose that  $(b^*, \lambda^*) \in B \times \Lambda$  is a voter equilibrium conditional on  $s^*$  and that  $s^*$  satisfies  $s_i^* \neq s_j^*$  for  $i \neq j$ . Also suppose that for exactly one  $j_1 \in K$ ,  $\hat{v}_{j_1}(s^*) > \hat{v}_{j}(s^*)$  for  $j \neq j_1$  and for exactly one  $j_2 \in K \setminus j_1$ ,  $\hat{v}_{j_2}(s^*) > \hat{v}_{j}(s^*)$  for  $j \neq j_1, j_2$ . Then for all  $j \neq j_1, j_2$  and  $j \in K$ ,  $v_{j}(b^*) = 0$ .

Proof: Suppose that  $(b^*, \lambda^*)$  is a voter equilibrium conditional on  $s^*$  satisfying  $s_i^* \neq s_j^*$  for  $i \neq j$ , such that  $v_i(b^*) > 0$  for some  $i \in K$  where  $i \neq j_1, j_2$ . It follows from the payoff function for informed voters that for all  $\alpha \in I$ ,  $b_{\alpha} \neq i$ . For uninformed voters, it follows from the proof of Theorem 4.1 that for  $\alpha \in U$  and for all  $j \in K$ ,

$$v_{\{1,\dots,j-1\}}(b^*) < F(y^*_{\alpha}) < v_{\{1,\dots,j\}}(b^*) \Longrightarrow b^*_{\alpha} = j,$$

and, defining  $x_j = F^{-1}(v_{\{1,...,j\}}(b^*))$  for  $j \in K$  and  $x_0 = F^{-1}(0)$ ,

$$x_{j-1} < y_{\alpha}^* < x_j \Longrightarrow b_{\alpha}^* = j.$$

This condition, which holds for all  $j \in K$ , implies that

$$y_{\alpha}^{*} < x_{j-1} \Longrightarrow b_{\alpha}^{*} \neq j$$
  
 $y_{\alpha}^{*} > x_{j} \Longrightarrow b_{\alpha}^{*} \neq j$ 

holds almost everywhere. So the votes for candidate i are coming from uninformed voters with ideal points satisfying  $x_{j-1} < y_{\alpha}^* < x_j$  and possibly a set of measure zero outside of this region. In other words,  $v_i(b^*) = F_U(x_i) - F_U(x_{i-1})$ .

As the distribution function  $F_I$  is strictly increasing, there is some  $\varepsilon > 0$  such that  $F_I(x_i) - F_I(x_{i-1}) > \varepsilon$ . But the definition of  $v_{\{1,\dots,i\}}(b^*)$  states that

$$v_{i}(b^{*}) = v_{\{1,\dots,i\}}(b^{*}) - v_{\{1,\dots,i-1\}}(b^{*})$$

$$= F(x_{i}) - F(x_{i-1})$$

$$= F_{U}(x_{i}) + F_{I}(x_{i}) - F_{U}(x_{i-1}) - F_{I}(x_{i-1})$$

$$> F_{U}(x_{i}) - F_{U}(x_{i-1}) + \varepsilon.$$

This is a contradiction with  $v_i(b^*) = F_U(x_i) - F_U(x_{i-1})$ . Therefore, it must be the case that  $v_i(b^*) = 0$ .

It is of course true that the informed voters are now endowed with a very high level of information about the electorate (as well as the candidates). Partly this is a reflection on the reasonableness of the assumptions imposed in the earlier sections. Indeed, this "high level" of information is usually assumed to be present among all voters in the usual models of strategic voting (Cox 1987; Palfrey 1989; Myerson and Weber 1993). Laying this response aside, it is remarkable that even if only a tiny fraction of voters have this high level of information, in equilibrium all voters act as if they were fully informed strategic voters. In this model, a small minority of

knowledgeable strategic voters can go a long way!

## 4.6 Conclusion

In this chapter, we have analyzed the relationship between voter information and election outcomes in a multicandidate setting. In our model with informed and uninformed voters, members of the latter group are able to make plausible inferences about the candidates' positions based on the vote share each candidate receives. In equilibrium, voters vote optimally given their beliefs, and beliefs are self-fulfilling in the sense that they are not contradicted by observable information. We have derived several interesting results. First, in the unique voter equilibrium of the model, all voters, informed and uninformed alike, vote as if they had perfect information. Second, we have described a dynamic process involving a sequence of polls that illustrates that this equilibrium is always reached. In addition, we have obtained results about candidate positioning equilibria when candidates are also uncertain about the characteristics of the voters. Finally, we have shown that if a small minority of voters are fully informed and use this information to vote strategically, in equilibrium all voters, including uninformed sincere voters, act as if they were voting strategically based on full information. The uninformed voters view the lack of support for trailing candidates by informed voters as evidence that these candidates are undesirable and react by voting for a more prominent candidate.

Of course, much work remains to be done in this area. For example, the implications of relaxing the assumptions in the model should be examined and the testable predictions of the model should be investigated with data from recent elections. This additional work should expand and strengthen the results presented here.

## Bibliography

- Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response," American Political Science Review, 69, 1218–1231.
- Argersinger, Peter H. 1980. "'A Place on the Ballot': Fusion Politics and Antifusion Laws," The American Historical Review, 85, 287–306.
- Bartels, Larry M. 1988. Presidential Primaries and the Dynamics of Public Choice. Princeton: Princeton University Press.
- Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee 1954. Voting: A Study of Opinion Formation in a Presidential Campaign. Chicago: University of Chicago Press.
- Black, Duncan 1958. The Theory of Committees and Elections. Cambridge: Cambridge University Press.
- Black, Jerome H. 1978. "The Multicandidate Calculus of Voting: Application to Canadian Federal Elections," American Journal of Political Science, 22, 609–638.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes 1960. The American Voter. Chicago: University of Chicago Press.
- Chung, Kai Lai 1974. A Course in Probability Theory, 2d. ed. Number 21 in Probability and Mathematical Statistics. Orlando: Academic Press.
- Converse, Philip E. 1964. "The Stability of Belief Elements Over Time." In Apter, David E., ed., *Ideology and Discontent*. New York: Free Press. Reprinted, in

- part, as Chapter 5 of Richard G. Niemi and Herbert F. Weisberg, eds., *Classics in Voting Behavior* (Washington D.C.: Congressional Quarterly Press, 1993).
- Converse, Philip E. and Gregory B. Markus 1979. "'Plus ça Change ...': The New CPS Election Study Panel," American Political Science Review, 73, 32–49. Reprinted, in part, as Chapter 8 of Richard G. Niemi and Herbert F. Weisberg, eds., Classics in Voting Behavior (Washington D.C.: Congressional Quarterly Press, 1993).
- Cox, Gary W. 1987a. "Duverger's Law and Strategic Voting," Mimeo. University of Texas.
- Cox, Gary W. 1987b. "Electoral Equilibrium under Alternative Voting Institutions," American Journal of Political Science, 31, 82–108.
- Cox, Gary W. 1991. "Duverger's Law: An Interpretive Review of Research," University of California, San Diego.
- Cox, Gary W. 1994. "Strategic Voting Equilibria Under the Single Non-Transferable Vote," American Political Science Review, 88, 608–621.
- Cukierman, Alex 1991. "Asymmetric Information and the Electoral Momentum of Public Opinion Polls," *Public Choice*, 70, 181–213.
- Dalton, Russel J. and Martin P. Wattenberg 1993. "The Not So Simple Act of Voting." In Finifter, Ada W., ed., *Political Science: The State of the Discipline II*, pages 193–218. Washington D.C.: American Political Science Association.
- Dharmadhikari, S. and K. Joag-dev 1988. *Unimodality, Convexity, and Applications*. San Diego: Academic Press.
- Downs, Anthony 1957. An Economic Theory of Democracy. New York: Harper and Row.
- Dudewicz, E. and S. Mishra 1988. *Modern Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons.

- Duverger, Maurice 1954. Political Parties: Their Organization and Activity in the Modern State. New York: Wiley. Translated by Barbara and Robert North.
- Duverger, Maurice 1986. "Duverger's Law: Forty Years Later." In Grofman, Bernard and Arend Lijphart, eds., *Electoral Laws and Their Political Consequences*, pages 69–84. New York: Agathon Press.
- Eaton, B. C. and R. Lipsey 1975. "The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition," Review of Economic Studies, 42, 27–49.
- Enelow, James M. and Melvin J. Hinich 1984. The Spatial Theory of Voting: An Introduction. Cambridge: Cambridge University Press.
- Feddersen, Timothy J. 1992. "A Voting Model Implying Duverger's Law and Positive Turnout," American Journal of Political Science, 36, 938–962.
- Ferejohn, J. and R. Noll 1988. "Three's a Crowd: Duverger's Law Reconsidered," Stanford University.
- Forsythe, Robert, Roger B. Myerson, Thomas A. Rietz, and Robert J. Weber 1993. "An Experiment on Coordination in Multi-candidate Elections: The Importance of Polls and Election Histories," *Social Choice and Welfare*, 10, 223–247.
- Gollin, Albert E. 1980. "Exploring the Liaison Between Polling and the Press," *Public Opinion Quarterly*, 44, 445–461.
- Greenberg, Joseph and Kenneth A. Shepsle 1987. "The Effects of Electoral Rewards in Multiparty Competition with Entry," *American Political Science Review*, 81, 525–537.
- Gutowski, William E. and John P. Georges 1993. "Optimal Sophisticated Voting Strategies in Single Ballot Elections Involving Three Candidates," *Public Choice*, 77, 225–247.
- Harsanyi, John C. 1967-68. "Games of Incomplete Information Played by Bayesian Players, Parts (I, II, and III)," *Management Science*, 14, 159–182, 320–334, 486–502.

- Hotelling, H. 1929. "Stability in Competition," Economic Journal, 39, 41–57.
- Johnson, N. and S. Kotz 1969. *Discrete Distributions*, volume 1. New York: Houghton Mifflin Company.
- Ledyard, John O. 1984. "The Pure Theory of Large Two Candidate Elections," *Public Choice*, 44, 7–41.
- Lian, Peng 1992. "Sequential Equilbria in Three-Candidate Elections," Caltech.
- McKelvey, Richard D. and Richard G. Niemi 1978. "A Multistage Game Representation of Sophisticated Voting for Binary Procedures," *Journal of Economic Theory*, 18, 1–22.
- McKelvey, Richard D. and Peter C. Ordeshook 1972. "A General Theory of the Calculus of Voting." In Herndon, J. F. and J.L. Bernd, eds., *Mathematical Applications in Political Science*, volume 6. Charlottesville: University Press of Virginia.
- McKelvey, Richard D. and Peter C. Ordeshook 1984. "Rational Expectations in Elections: Some Experimental Results Based on a Multidimensional Model," *Public Choice*, 44, 61–102.
- McKelvey, Richard D. and Peter C. Ordeshook 1985a. "Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources," *Journal of Economic Theory*, 36, 55–85.
- McKelvey, Richard D. and Peter C. Ordeshook 1985b. "Sequential Elections with Limited Information," American Journal of Political Science, 29, 480–512.
- McKelvey, Richard D. and Peter C. Ordeshook 1986a. "Information, Electoral Equilibria, and the Democratic Ideal," *Journal of Politics*, 48, 909–937.
- McKelvey, Richard D. and Peter C. Ordeshook 1986b. "Sequential Elections with Limited Information: A Formal Analysis," *Social Choice and Welfare*, 3, 199–211.

- McKelvey, Richard D. and Peter C. Ordeshook 1987. "Elections with Limited Information: A Multidimensional Model," *Mathematical Social Sciences*, 14, 77–99.
- McKelvey, Richard D. and Peter C. Ordeshook 1990a. "A Decade of Experimental Research on Spatial Models of Elections and Committees." In Enelow, James M. and Melvin J. Hinich, eds., *Advances in the Spatial Theory of Voting*, pages 99–144. Cambridge: Cambridge University Press.
- McKelvey, Richard D. and Peter C. Ordeshook 1990b. "Information and Elections: Retrospective Voting and Rational Expectations." In Ferejohn, John and J. Kuklinski, eds., Information and Democratic Process, pages 281–312. Champaign: University of Illinois Press.
- Myerson, Roger B. and Robert J. Weber 1993. "A Theory of Voting Equilibria," American Political Science Review, 87, 102–114.
- Nie, Norman H. and Kristi Andersen 1974. "Mass Belief Systems Revisted: Political Change and Attitude Structure," *Journal of Politics*, 36, 665–678. Reprinted, in part, as Chapter 6 of Richard G. Niemi and Herbert F. Weisberg, eds., *Classics in Voting Behavior* (Washington D.C.: Congressional Quarterly Press, 1993).
- Nie, Norman H., Sydney Verba, and John R. Petrocik 1976. The Changing American Voter. Cambridge: Harvard University Press.
- Niemi, Richard G. and Herbert F. Weisberg, eds. 1993. Classics in Voting Behavior. Washington D.C.: Congressional Quarterly Press.
- Ordeshook, Peter C. and Thomas R. Palfrey 1988. "Agendas, Strategic Voting, and Signaling with Incomplete Information," American Journal of Political Science, 32, 441–466.
- Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry," Review of Economic Studies, 51, 139–156.
- Palfrey, Thomas R. 1989. "A Mathematical Proof of Duverger's Law." In Ordeshook, Peter C., ed., *Models of Strategic Choice in Politics*, pages 69–92: University of Michigan Press.

- Palfrey, Thomas R. and Howard Rosenthal 1991. "Testing Game-Theoretic Models of Free Riding: New Evidence on Probability Bias and Learning." In Palfrey, Thomas R., ed., *Laboratory Research in Political Economy*, pages 239–268: University of Michigan Press.
- Popkin, Samuel L. 1991. The Reasoning Voter. Chicago: University of Chicago Press.
- Popkin, Samuel L., John Gorman, Jeffrey Smith, and Charles Phillips 1976. "What Have You Done for Me Lately: Toward an Investment Theory of Voting," American Political Science Review, 70, 779–805.
- Rae, Douglas 1967. The Political Consequences of Electoral Laws. New Haven: Yale University Press.
- Ratzan, Scott C. 1989. "The Real Agenda Setters: Pollsters in the 1988 Presidential Campaign," American Behavioral Scientist, 32, 451–463.
- Riker, William H. 1976. "The Number of Political Parties: A Reexamination of Duverger's Law," Comparative Politics, 9, 93–106.
- Riker, William H. 1982. "The Two-party System and Duverger's Law: An Essay on the History of Political Science," *American Political Science Review*, 76, 753–766.
- Riker, William H. 1986. "Duverger's Law Revisited." In Grofman, Bernard and Arend Lijphart, eds., *Electoral Laws and Their Political Consequences*, pages 19–42. New York: Agathon Press.
- Schattschneider, E. E. 1942. Party Government. New York: Farrar and Rinehart.
- Shepsle, K. 1991. Models of Multiparty Electoral Competition, volume 45. Chur, Switzerland: Harwood Academic Publishers.
- Simon, Herbert A. 1954. "Bandwagon and Underdog Effects of Election Predictions," Public Opinion Quarterly, 18.
- Smith, Eric R.A.N. 1989. The Unchanging American Voter. Berkeley: University of California Press.

- Smith, Tom W. 1990. "The First Straw? A Study of the Origins of Election Polls," *Public Opinion Quarterly*, 54, 21–36.
- Weber, Shlomo 1992. "On Hierarchical Spatial Competition," Review of Economic Studies, 59, 407–425.
- West, Darrell M. 1991. "Polling Effects in Election Campaigns," *Political Behavior*, 13, 151–163.
- Wright, Stephen G. and William H. Riker 1989. "Plurality and Runoff Systems and Numbers of Candidates," *Public Choice*, 60, 155–175.