# SOME MULTIVARIATE PROBLEMS
# OF A SPATIAL MODEL OF VOTING
# UNDER MAJORITY RULE

by

R. W. Hoyer

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Dr. Lawrence S. Mayer
Co-chairman
Department of Statistics

Dr. Raymond H. Myers
Co-chairman
Department of Statistics

Dr. Jesse C. Arnold
Department of Statistics

Dr. Boyd Harshbarger
Department of Statistics

Dr. James Herndon
Department of Political Science

Dr. Leon W. Rutland
Department of Mathematics

Dr. Robert M. Todd
Department of Mathematics

October 1976

Blacksburg, Virginia

## ACKNOWLEDGEMENTS

The author wishes to express profound gratitude to Lawrence S. Mayer, without whose inspiration, encouragement, enthusiasm, and friendship this dissertation would not have been possible.

The author's Graduate Committee made many contributions to the research reported in this dissertation, and individual members of that Committee were most helpful at such time as university bureaucratic structures seemed to be designed to impede--rather than enhance--his graduate progress.

Judith Youngblood Hoyer deserves much praise for her patient encouragement and support throughout many years of the author's graduate study.

There are plaudits for Ruth Zemel whose cooperation, good humor, and skill in typing this manuscript made an otherwise unpleasant task quite bearable.

Finally, the author is deeply indebted to his parents, Robert B. Hoyer and Mary Reed Hoyer, who made many sacrifices which enabled him to obtain a Bachelor of Arts degree, and who encouraged him to pursue a program of graduate studies. This dissertation is dedicated to the memory of the author's father, Robert B. Hoyer.

# TABLE OF CONTENTS

Page

## LIST OF ILLUSTRATIONS

| Figure | | Page |
|---|---|---|

# CHAPTER I

## INTRODUCTION

This dissertation is concerned with a number of outstanding problems in the analysis of collective decision-making.[1] The essential question in this area of study is the following: Given a collection of individuals, each of whom has a preference ordering on a set of objects or social states, is it possible to "aggregate" those preferences to produce a "well-behaved" social preference ordering in a manner which is acceptable to a majority of the members of society? While descriptive, this is obviously an imprecise statement of the problem, so we will restate it in a more formal terminology in the next chapter.

The question and its answer are clearly of the utmost importance to both economists and political scientists since they are basically interested in studying aggregate choices and decisions. For example, the <u>majority rule decision function</u> has probably been the most widely utilized political collective-decision rule in Western civilization during the past two centuries. In the economic realm, the basis of the <u>market system</u> depends upon the producer's ability to interpret and reshape consumer demand, which is

[1] In the economics and political science literature, collective decision-making is alternately referred to as public choice, social choice, voting theory, and collective action.

ultimately an aggregate function of rather complex individual preference orderings. Strangely enough, while most citizens in democratic, capitalistic societies seem to have a reasonably accurate intuitive feeling for how these economic and political systems function, only a miniscule number of political scientists and economists is cognizant of the formal foundations of these systems. Furthermore, we conjecture that the vast majority of citizens in our society has almost unyielding confidence, however misplaced, in the internal consistency of the structure of their political and economic systems. It is our experience that while most individuals would readily agree that there are rather serious practical problems inherent to functioning as a capitalistic democracy, they would just as quickly disagree with the assertion that there are logical inconsistencies underpinning such a societal structure. Yet, as we shall demonstrate in the second chapter, two of the components of the foundation of such a political-economic system (in particular, voting under majority rule and preference aggregation in a free market) have logical inconsistencies which are not easily circumvented.

It is by virtue of the fact that collective decision-making can be characterized in mathematical terms that the area is pregnant with problems of interest to mathematicians. Furthermore, since many models of public choice are stochastic, it is an area ripe with interesting problems for theoretical statisticians and probabilists. In addition, models of political and economic phenomena must be validated, so there is a wealth of interesting and important problems related to collective decision-making for applied statisticians and data analysts.

It is certainly the case that economic and political theorists have made, and will continue to make, important theoretical and substantive contributions to the understanding of public choice; nevertheless, perusal of the "mathematical quality" of this literature will reveal that, at the present stage in the development of models of collective choice, cooperative ventures by <u>applied mathematicians</u> in consort with economists and political scientists would be beneficial to all parties. Such an interaction would open important new areas of application for the mathematicians and provide the more mathematically sophisticated economists and political scientists with a type of formal guidance which would be very useful--and probably most fruitful.

It is beyond argument that the most critical shortcoming of the current theory of public choice is the fact that existing models have not been empirically tested. Until recently, this shortcoming was inconsequential, since existing models were so rudimentary that few theorists would have expected them to stand the test of empirical scrutiny.

Collective decision-making models, however, are becoming increasingly more refined, and since empirical and mathematical validations are essential parts of the modeling process, it is important that collective-choice models be subjected to statistical analyses. While most political theorists recognize this problem, few are either interested in making or capable by training to make a conscientious effort to eliminate it, and few statisticians have become interested in such problems.

In Chapter II of this dissertation, we define the essential objective of collective decision-making in a mathematically precise manner. Arrow's General Impossibility Theorem is stated, and its practical ramifications for a system of collective action are discussed. A body of work--formulated within the context of a spatial theory--embraces all of the axiomatic structure of the Impossibility Theorem except the "Universal Domain" axiom, and one of its purposes is to determine conditions under which "well-behaved" social welfare functions obtain. The spatial model is developed in Chapter III, but it is restrictive in several important respects. In particular, although a provision has been made to allow citizens (consumers) to express their preferences over social (commodity) states in a completely individual manner, no provision has been made for them to express loss <u>individually</u> if they must settle for less than their most preferred outcomes. In addition, the model includes the assumption that citizens vote (and consumers make purchases) even when they find the attainable alternatives extremely distasteful. A generalized spatial model in which the first objection is eliminated is developed, and the conclusions under the extended model are contrasted with those of the more rudimentary spatial formulation. In Chapter IV, we show that the only published responses to the second deficiency of the spatial model do not generalize the model presented in Chapter III. In fact, there are subtle errors in those models which supposedly "allow" voters (and consumers) to abstain when they are alienated from alternative social (or commodity) states. Furthermore, we show that any effort to repair these errors leads to a model which is unacceptable because it includes interpersonal comparisons of utility. Consequently, although we have not formulated a model which includes abstentions due to alienation, we have shown that existing models which purport to do so are not mathematically sound. We, therefore, reopen an important area of study for those interested in spatial formulations of political or economic behavior.

In Chapter V, we compare sets of equilibrium strategies for candidates for office when the candidates are attempting to maximize certain objective functions, and, to be specific, we show that the set of equilibria is invariant under seven important objective functions. The proposed model, while not restricted to comparing election outcomes in a spatial setting, is completely compatible with the spatial formulation. Furthermore, the model includes a provision to account for the fact that some citizens may abstain from voting.

The final chapter contains concluding remarks, including a brief assessment of our opinion regarding the utility of spatial models for the analysis of political and economic phenomena.

# CHAPTER II

## BACKGROUND

The mathematical theory of election by "majority rule" has a long, if sporadic, history. In the latter half of the eighteenth century, Borda, Condorcet, and Laplace each made independent attempts to construct a mathematical model for voting in a democracy. In the last decade of that century, E. J. Nanson re-examined the results of these mathematicians and devised several additional systems for voting under somewhat restrictive conditions. Slightly less than a century later, the Rev. C. L. Dodgson (Lewis Carroll) made important contributions to the theory and published his results in three pamphlets. In the 1950's, the proof of Arrow's famous Impossibility Theorem, the publication of Duncan Black's book, The Theory of Committees and Elections, and Anthony Downs' intuitive attempt to utilize economic models to explain the structure of party politics in a democratic society revitalized interest in and gave new direction to the problem of election by majority rule.[1] It was Arrow's work, however, which inspired the plethora of research activity during the past ten years--including much of the

[1]Kenneth Arrow, Social Choice and Individual Values, 2nd edition, (New York: John Wiley and Sons, 1963); Duncan Black, The Theory of Committees and Elections, (Cambridge: Cambridge University Press, 1958); Anthony Downs, Economic Theory of Democracy, (New York: Harper and Row Publishers, 1957).

analysis contained in this dissertation--directed at determining how "rational" social decisions should be made, based upon knowledge of individual preference orderings. We will set the stage for the remaining chapters of this dissertation by reviewing Arrow's Theorem.

## Section 2.1 The Fundamental Objective of Collective Decision-Making

James Buchanan and Gordon Tullock describe collective action as

> the action of individuals when they choose to accomplish purposes collectively rather than individually, and the government is seen as nothing more than the set of processes, the machine, which allows such collective action to take place.[2]

Kenneth Shepsle, elaborating on the observation of Buchanan and Tullock, wrote,

> Theories of collective choice, as I conceive them, concern the implementation and conduct of decision-making processes by and/or for collections of individuals, as well as the enforcement and administration of the decisions

[2] James Buchanan and Gordon Tullock, The Calculus of Consent, (Ann Arbor: The University of Michigan Press, 1967), p. 13.

that emerge from these processes. While various theories may alternately emphasize implementation (a theory of constitutions), conduct (a theory of institutions), or enforcement and administration (a theory of social control), most share the properties suggested above, namely:

(1) they tend to be general theories of collectivities,

(2) they rely on the individualistic perspective and the assumption of purposeful behavior, and

(3) they employ some form of the choice paradigm to link purpose with behavior.[3]

With no apologies for redundancy, we will restate what we believe is the essential objective of collective decision-making. To wit: given a collection of citizens, each of whom has an individual preference ordering on a set of social (commodity) states, is it possible to "aggregate" those preferences to produce a "well-behaved" social preference ordering in a manner which is acceptable to the members of society? It should be clear that there are seven concepts

[3] Kenneth Shepsle, "Theories of Collective Choice" in Cornelius Cotter, et al., eds., Political Science Annual, V: Collective Decision-Making, (Indianapolis: Bobbs-Merrill, 1973), pp. 1-87.

(those which have been underscored) in the above sentence which require formal definition and explanation. Certain of these concepts fall under the rubric of the individual voter's choice calculus, so we discuss those ideas first.

We assume that society is a class of individuals, and represent the generic $i^{th}$ voter (citizen) by $V_i$.[4] We will have much to say about the collection of such individuals later, but, for the moment, we will concentrate on $V_i$'s voting calculus.

First, there is a set $\Omega$ of social states (or commodities) over which $V_i$ has differential preference. Although, in many situations, the space of social states is a finite set, we have no objection to assuming that it contains an infinite number of elements. Indeed, in the spatial model developed in the next chapter, the commodity space is represented as an uncountably infinite subspace of Euclidean n-space $E^n$.

The individuals in this social structure do not choose between outcomes directly, because some social states, while highly desirable, may be virtually inaccessible. In essence, individuals act in a manner which they believe effects various outcomes, and they choose between actions so that their

---

[4] Of course, we could specify that these N individuals are "consumers" and couch our terminology in the language of economists. For the most part, we will describe the models in this dissertation from a political perspective.

preferred social states eventuate as a result of those actions. In short, there is a set $A$ of actions and, for the $i^{th}$ voter, a function $f_i: A \to \Omega$ which embodies $V_i$'s subjective, causal view of the world. If $f_i$ is deterministic, then $f_i(a_0) = o_0$ means that the $i^{th}$ voter believes that if he takes action $a_0$, the outcome $o_0$ will ensue.[5]

The second component of the individual's voting calculus is his preference ordering on the social states in $\Omega$. In particular, we assume that $V_i$ is <u>rational</u> in the sense that he does have a subjective, causal view of the world and in the sense that he has a complete, transitive, binary preference ordering $R_i$ on the elements in $\Omega$. Thus, for each voter $V_i$ there exists an ordering $R_i$ (that is, a function from $\Omega \times \Omega$ into $\Omega$) such that

i). if $o_1, o_2 \varepsilon \Omega$, then either $o_1 R_i o_2$ or $o_2 R_i o_1$.

ii). if $o_1, o_2, o_3 \varepsilon \Omega$, and $o_1 R_i o_2$ and $o_2 R_i o_3$, then $o_1 R_i o_3$.

If $o_1 R_i o_2$, then <u>we say $V_i$ prefers $o_1$ at least as much as $o_2$</u>. If $o_1 R_i o_2$ and it is not the case that $o_2 R_i o_1$, then the $i^{th}$

[5]Of course, $f_i$ may not be deterministic. The $i^{th}$ voter $V_i$ may believe that if he takes action $a_0$, then any one of the outcomes in $\Omega$ may eventuate, with $o_k$ occurring with probability $p_k$. In this case, the image under $f_i$ of an action $a_0$ is a lottery on the outcomes in $\Omega$, and $V_i$'s preference ordering on the set of actions will be defined in terms of something on the order of expected utility.

voter strictly prefers $o_1$ to $o_2$, and we represent that with the shorthand notation $o_1 P_i o_2$. If $o_1 R_i o_2$ and $o_2 R_i o_1$, the $i^{th}$ voter is indifferent between those two social states, and we abbreviate $o_1 I_i o_2$.

The individual's causal view of the world $f_i$, which links his possible actions with the various accessible social states, along with his preference ordering $R_i$ on those social states, induces a preference ordering $R_i'$ on the elements in $A$ in the following manner: if $a_1$, $a_2 \varepsilon A$, then $a_1 R_i' a_2$ if and only if $f_i(a_1) R_i f_i(a_2)$. Of course, $P_i'$ and $I_i'$ can be defined on the action space analogously. Even though the spatial model to be developed in the next chapter could be formulated with fundamental elements taken from the voters' action space $A$, we choose to take the social states in $\Omega$ as the basic objects in the model. The principal justification for this choice is the fact that virtually no research has been conducted with regard to the nature of the functions $f_i$, so considerable uncertainty would be introduced in a model based upon the action space $A$.

Given a particular choice situation, the $i^{th}$ voter has a <u>choice</u> (or maximal) <u>set</u> $C_i(\Omega) = \{o \mid \text{if } o' \varepsilon \Omega, \text{ then } o R_i o'\}$. We assume that the individual makes his choice by selecting an element $o \varepsilon C_i(\Omega)$.

Having outlined the basic components of the individual's choice calculus, we turn to the fundamental characteristics of a social choice calculus. First, we assume the existence of a finite collection $V = \{V_1, V_2, \ldots, V_N\}$ of N voters, where $N \geq 3$. We see no advantage to expanding the collection of individuals to an infinite set; however, there are both political and economic theorists who maintain that such a generalization is desirable.[6]

Now, let $\mathcal{R}$ be the collection of all complete, transitive, binary relations on $\Omega$. The voters' preferences over the space of social states is an N-tuple, $(R_1, R_2, \ldots, R_N) \varepsilon \mathcal{R}^N$, and the "aggregation" formula is a function $F: \mathcal{R}^N \rightarrow \mathcal{R}$. If $F(R_1, R_2, \ldots, R_N) = R$, then the social preference ordering R represents "society's choice" among the elements of $\Omega$ under the system associated with F. Since $R \varepsilon \mathcal{R}$, the system is "well-behaved" in the sense that it is complete (it orders all of the social states in $\Omega$) and transitive.

The function F which "aggregates" individual preferences and generates a social preference ordering is called a social decision function or a social welfare function. To complete the formalization of the essential objective of collective decision-making, it remains to explain what "a manner which is acceptable to the members of society"

---

[6] Richard McKelvey, "Policy Related Voting and Electoral Equilibrium," Econometrica, Vol. 43 (1975), pp. 815-843. For an interesting criticism of this viewpoint, see Philip Miller, "Criticisms of McKelvey's Theory of Policy Related Voting," (unpublished manuscript).

means. The usual definition of what is socially acceptable is the set of normative constraints on F formulated as axiomatic conditions for Arrow's theorem. In particular, they are

Condition I (Universal Domain): F must be a function whose domain is $\mathcal{R}^N$.

Condition II (Pareto Principle): If given $o_1, o_2 \varepsilon \Omega$ and $(R_1, R_2, \ldots, R_N) \varepsilon \mathcal{R}^N$ such that $o_1 R_i o_2$ for $i = 1, 2, \ldots, N$ and if $F(R_1, R_2, \ldots, R_N) = R$, then $o_1 R o_2$.

Condition III (Independence of Irrelevant Alternatives): Let $C(\Omega)$ be the choice set of $\Omega$, let $o \varepsilon \Omega - C(\Omega)$, and let $\Omega' = \Omega - \{o\}$. If we are given $(R_1, R_2, \ldots, R_N) \varepsilon \mathcal{R}^N$ and if $(R_1', R_2', \ldots, R_N')$ is an N-tuple of preference orderings on $\Omega'$ such that for $o_1, o_2 \varepsilon \Omega$ with

[7] For excellent formal discussions of Arrow's General Impossibility Theorem see James Quirk and Rubin Saposnik, Introduction to General Equilibrium Theory and Welfare Economics, (New York: McGraw Hill, 1968), pp. 103-124 or Shepsle's "Theories of Collective Choice."

$$o_1 \neq o \neq o_2,\ o_1 R_i o_2 \implies o_1 R_i' o_2$$
for $i = 1, 2, \ldots, N$; then
$C(\Omega) = C(\Omega')$.[8]

<u>Condition IV</u> (Non-dictatorship): If $(R_1, R_2, \ldots, R_N) \varepsilon \mathcal{R}^N$ and $F(R_1, R_2, \ldots, R_N) = R$, then there exists no $i \varepsilon \{1, 2, \ldots, N\}$ such that whenever $o_1, o_2 \varepsilon \Omega$, $o_1 R_i o_2 \iff o_1 R o_2$.

The first condition simply stipulates that, regardless of the existing combination of N individual preference orderings, the social welfare function <u>must</u> generate a social preference ordering.

If the Pareto Principle holds, then, at least in one sense, the social preference ordering cannot be perverse. In particular, if every voter in the social order prefers the social state $o_1$ at least as much as the social state $o_2$, then the social preference ordering must prefer $o_1$ to $o_2$. Second, this condition stipulates that the social choice cannot be imposed on the electorate from outside the social

[8]The social choice set is defined in a manner analogous to the definition of the individual's choice set, that is, if R is the social preference ordering on the space of social states $\Omega$, then $C(\Omega) = \{o \mid \text{if } o' \varepsilon \Omega, \text{ then } oRo'\}$.

structure; for if that were possible, then by such an imposition it may be that $o_2Ro_1$ even if for every voter, $o_1R_io_2$.

The Independence of Irrelevant Alternatives is perhaps the most controversial of Arrow's conditions. To understand it, we assume that $o \epsilon \Omega$ is not in the social choice set $C(\Omega)$, and we assume that the social welfare function $F$ maps the individual orderings into a social ordering $R$ such that, as far as the social states $o_1$ and $o_2$ are concerned, $o_1Ro_2$. Now, if, after the irrelevant alternative $o$ is eliminated from consideration, each voter ranks $o_1$ at least as high as he did initially relative to $o_2$, and if the new social ordering is $R'$, then it must follow that $o_1R'o_2$.

The final condition stipulates that there can be no individual (dictator) in the social order such that whatever that individual's preference ordering on the social states is, it necessarily follows that the social preference ordering will be identical to it.

## Section 2.2 Arrow's Theorem and Its Importance

The background material in the previous section, in addition to specifying both the individual citizen's voting calculus and the societal voting calculus, also sets the stage for Arrow's Impossibility Theorem.

Theorem 2.1 (Arrow): Given a set of social states $\Omega$ containing at least three elements, and given an N-tuple $(R_1, R_2, \ldots, R_N) \varepsilon \mathcal{R}^N$ of individual preference orderings on $\Omega$, where $N \geq 3$, then there exists no social welfare function $F: \mathcal{R}^N \to \mathcal{R}$ satisfying Conditions I - IV.

In essence, Arrow's Theorem specifies that if we are committed to having a social welfare function generate a social preference ordering, then we cannot expect that function to satisfy the normative constraints set forth by Arrow. Conversely, if the normative conditions hold, and if $F$ is a function (not a social welfare function) from $\mathcal{R}^N$ into an appropriate range space $\mathcal{R}'$, then $F(R_1, R_2, \ldots, R_N)$ will not in general be a binary ordering--it will either fail to be complete or it will not be transitive.

It is difficult to measure the impact of the General Impossibility Theorem on certain segments of the economics and political science communities. It simultaneously provided new insight into and inspired some pessimism about the structure of the democratic process since one of the cornerstones of that process--namely, voting under majority rule--was seen to be simply another normatively defective social welfare function.[9] In any event, numerous attempts

[9] An equally severe blow was dealt to preconceptions of the economic system since preference aggregation in a free market was also seen to have structural deficiencies.

were made to rationalize the difficulties made explicit by Arrow's paradox. On another plane, efforts to reconcile the paradox with the unyielding faith of political theorists in the foundation of the democratic process led to many experiments with the formal structure of that process. One such experiment attempts to identify the weakest possible constraints on the preference orderings of individuals which guarantees the existence of an otherwise "well-behaved" social welfare function. The spatial formulation developed in the next chapter is such a model, that is, one in which restrictive assumptions regarding the "Universal Domain" condition are specified in order to insure a complete, transitive social preference ordering in the range space of a social welfare function satisfying Conditions II - IV.

# CHAPTER III

## A SPATIAL MODEL OF VOTING UNDER MAJORITY RULE

During the score of years since Arrow proved his famous Impossibility Theorem, political and economic theorists have endeavored to understand and explain social structures generated by aggregating citizens' preference orderings with a social welfare function such as majority rule. An explanatory model of considerable interest and one demonstrating a reasonable degree of mathematical sophistication is a spatial formulation of voting under majority rule which was conceived by Otto Davis and Melvin Hinich about ten years ago and refined by a number of interested parties during the past decade.[1]

A model of any political phenomena is, by its very nature, an oversimplification of some condition or process; so it invariably contains descriptive flaws. The most apparent shortcomings of the Davis-Hinich model (henceforth abbreviated the D-H model) are

1. no _formal_ provision is made to allow different voters to experience different modes or magnitudes of "unhappiness" if a less than ideal social state is imposed upon them.

[1] Otto Davis and Melvin Hinich, "A Mathematical Model of Policy Formation in a Democratic Society," in Joseph Bernd ed., _Mathematical Applications in Political Science II_ (Dallas: Southern Methodist University Press, 1966), pp. 175-208.

2. no formal provision is made to allow a citizen to abstain from voting if he is sufficiently unenamored with the social states adopted as political platforms by the vying candidates.

Details of a model more general than the D-H model will be put forth in this chapter, and the principal advantage of the abstraction is that the new model is a more faithful representation of the "real world" in the sense that differential loss functions are built into its structure. Thus, the first deficiency of the D-H model will be eliminated.

## Section 3.1 Preliminaries

In this particular model we assume that there is a collection $V = \{V_1, V_2, \ldots, V_N\}$ of N citizens, none of whom may abstain from voting, and a set of p candidates $\{C_1, C_2, \ldots, C_p\}$ who are contending for a certain office. We are able to identify the salient campaign issues and describe them in such a way that a voter can indicate his most preferred position on a specific issue by reporting a value of some numerical index.[2] Hence, the $i^{th}$ voter's

[2] In fact, we interpret campaign issues in a rather broad sense. For example, one numerical index may be the weight (as indicated on a "thermometer scale") which the citizen assigns to a specific policy decision - such as the importance of raising taxes in order to finance a specific social welfare program. Another index might be the percentage of each tax dollar which the citizen would like to see allocated for defense expenditures. Still another possibly relevant numerical index is the voter's evaluation of the importance of filling the next vacancy on the Supreme Court with a female Justice.

preference position on n issues of policy is a vector $x_1 = [x_{i1}, x_{i2}, \ldots, x_{in}]'$ in n-dimensional Euclidean space $E^n$, where $x_{ik}$ is the value assigned by the $i^{th}$ voter to the $k^{th}$ issue. Similarly, the $j^{th}$ candidate's political platform can be represented by the vector $\theta_j = [\theta_{j1}, \theta_{j2}, \ldots, \theta_{jn}]'$ in $E^n$.[3] Furthermore, the $i^{th}$ voter experiences a non-negative loss with respect to the election of the $j^{th}$ candidate defined by the function

$$L_i(\theta_j) = \psi_i[(x_i - \theta_j)'A_i(x_i - \theta_j)], \tag{3.0}$$

where $A_i$ is an $n \times n$ symmetric, positive definite matrix and $\psi_i[\cdot]$ is a monotone increasing function. We frequently refer to the utility $U_i(\theta_j)$ that the $i^{th}$ voter receives if the $j^{th}$ candidate is elected instead of the corresponding loss. Presumably $V_i$ will receive some maximum utility $\lambda_i$ if the candidate's strategy coincides with his most preferred point, and his utility will diminish as the candidate's platform and his preferred point become more disparate. A class of utility functions consistent with these restrictions is defined by

$$U_i(\theta_j) = \lambda_i - \psi_i[(x_i - \theta_j)'A_i(x_i - \theta_j)], \tag{3.1}$$

where $\psi_i$ is the function specified in (3.0). Since, for each citizen, $\lambda_i$ is a constant, it is a trivial matter to

---

[3]It is assumed that the candidate's precise position in $E^n$ is known by all citizens. This is one of those oversimplifications that we are willing to tolerate for the time being.

construct functions $h_i[\cdot]$ for $i = 1,\ldots, N$ such that

(3.2) $$U_i(\theta_j) = h_i[(x_i - \theta_j)'A_i(x_i - \theta_j)],$$

where $h_i$ is a monotone decreasing function (see Figure 3.1). Structurally, each citizen's utility function is a mapping from $E^n$ into $E^1$ satisfying the condition that as his loss $L_i(\theta_j)$ increases, his utility $U_i(\theta_j)$ decreases. In short, we may use loss functions and utility functions interchangeably since both are monotone functions of the same argument provided that we are careful to recognize that their directions of monotonicity differ.

It is interesting to note how $V_i$'s utility with regard to various social states is related to his preference over those social states. That relationship is specified according to the following definition.

Definition 3.1: If $\theta_1, \theta_2 \in E^n$, then the $i^{th}$ voter is said to prefer $\theta_1$ at least as much as $\theta_2$, and we write $\theta_1 R_i \theta_2$, if $U_i(\theta_1) \geq U_i(\theta_2)$. If $U_i(\theta_1) > U_i(\theta_2)$, then $V_i$ strictly prefers $\theta_1$ to $\theta_2$ and we write $\theta_1 P_i \theta_2$. If $U_i(\theta_1) = U_i(\theta_2)$, we say he is indifferent between $\theta_1$ and $\theta_2$ and abbreviate $\theta_1 I_i \theta_2$.

This definition is slightly misleading since it would appear to treat utility as the prior concept when quite the opposite is the case, that is, if $V_i$ prefers $\theta_1$ over $\theta_2$, then, if he is rational in a formal sense, it should follow

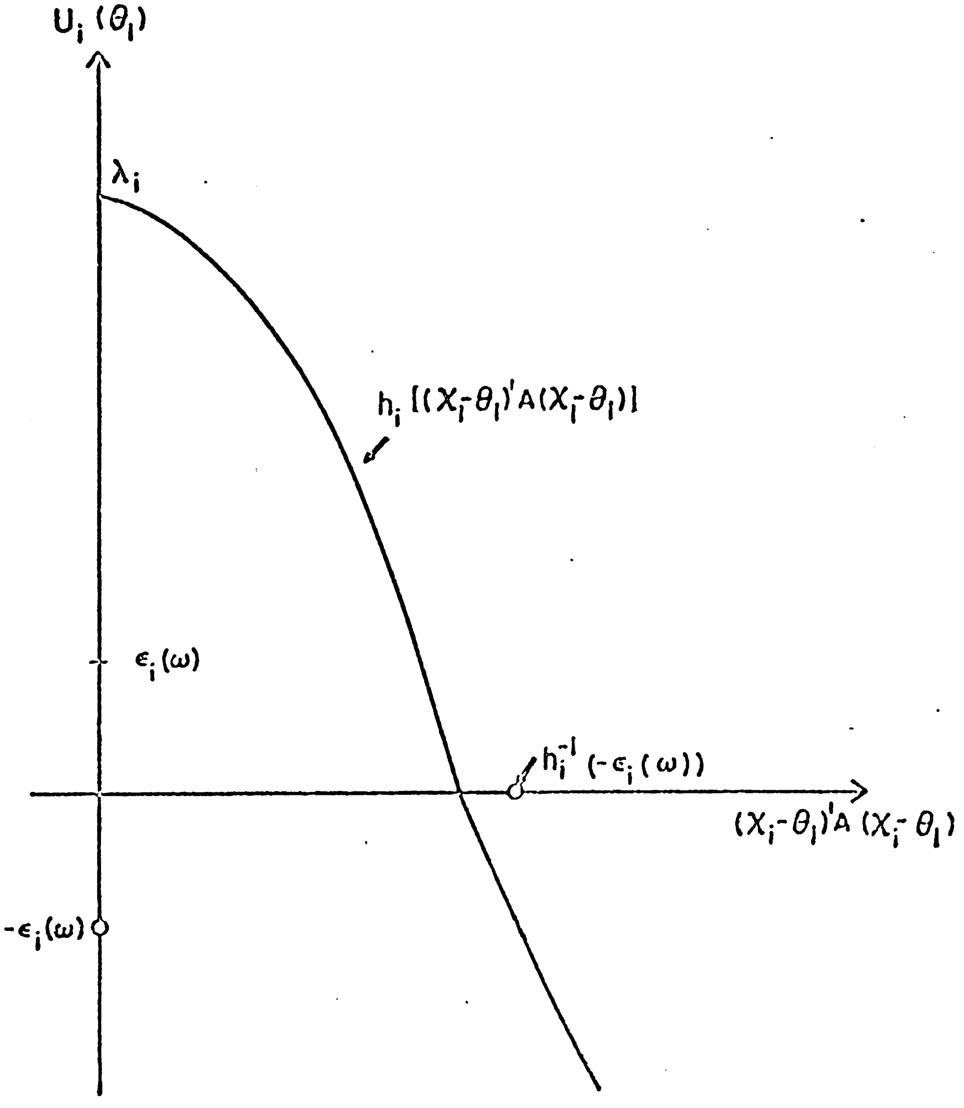



Figure 3.1

The $i^{th}$ Voter's Utility Defined as a Function of the Normed Distance from His Preferred Point to the Candidate's Strategy Vector that $U_i(\theta_1) > U_i(\theta_2)$. In short, preference is the prior concept, and utilities of social states are ordered--if not completely determined--according to the $i^{th}$ citizen's preference ordering on the space of social states.

There are two important observations which must be made concerning $V_i$'s utility function. First, we note that, although the $i^{th}$ citizen's preference ordering $R_i$ is related to his utility function $U_i$ by Definition 3.1, it is not essential that we know the function $h_i$ (see equation (3.2)) in order to specify $R_i$. To see this, suppose that $U_i(\theta_1) \geq U_i(\theta_2)$ holds for the $i^{th}$ voter whose loss matrix is $A_i$. By definition, $U_i(\theta_1) \geq U_i(\theta_2)$ is equivalent to $h_i[(x_i - \theta_1)'A_i(x_i - \theta_1)] \geq h_i[(x_i - \theta_2)'A_i(x_i - \theta_2)]$. But because of the monotonicity of $h_i$, the above inequality is true if and only if $(x_i - \theta_1)'A_i(x_i - \theta_1) \leq (x_i - \theta_2)'A_i(x_i - \theta_2)$. In other words, a voter who is attempting to make a choice between two candidates $C_1$ and $C_2$, whose platforms are $\theta_1$ and $\theta_2$ respectively, based upon a comparison of $U_i(\theta_1)$ and $U_i(\theta_2)$ would make precisely the same decision by comparing the two normed distances $||x_i - \theta_1||_{A_i}$ and $||x_i - \theta_2||_{A_i}$, where $V_i$'s norm is defined by

(3.3) $$||x_i - y||^2_{A_i} = (x_i - y)'A_i(x_i - y) \quad \text{for all } y \in E^n.$$

While our initial inclination is to characterize the $i^{th}$ voter by the ordered triple $(x_i, A_i, h_i)$, where $x_i$ is his most preferred point, $A_i$ is his loss matrix, and $h_i$ is the monotone decreasing function in (3.2), since we are not interested in the $i^{th}$ citizen's utility per se, but only as it relates to his preference ordering $R_i$, it will suffice to characterize the $i^{th}$ voter by the ordered pair $(x_i, A_i)$, that is, the function $h_i$ is excess baggage and may be discarded. This is one of the exciting characteristics of both the original D-H model and our more general formulation, because it allows the political scientist or economist to construct individual preference orderings, aggregate them in some manner, and, all the while, avoid making interpersonal comparisons of utility.[4]

In $E^n$, the matrix $A_i$ determines n-dimensional ellipsoids of constant loss which are "centered" at the $i^{th}$ voter's most preferred point.[5] Figure 3.2 depicts a two-dimensional issue space $E^2$ with three voters, $V_1$, $V_2$, and $V_3$, whose most preferred points and indifference ellipses are indicated and where two candidates' political platforms

---

[4]The $i^{th}$ voter's preference ordering $R_i$ can be defined in terms of his norm $||\cdot||_{A_i}$ on the space of social states in the following manner: if $y, z \varepsilon E^n$, then the $i^{th}$ voter $(x_i, A_i)$ prefers $y$ at least as much as $z$, and we write $yR_iz$, if and only if $||x_i - y||_{A_i} \leq ||x_i - z||_{A_i}$. Of course, $P_i$ and $I_i$ can be similarly defined.

[5]By an ellipsoid of constant loss we mean that if $(x_i - \theta_1)'A_i(x_i - \theta_1) = (x_i - \theta_2)'A_i(x_i - \theta_2)$, then $\theta_1$ and $\theta_2$ are on the same ellipsoid with center at $x_i$, and the $i^{th}$ voter assigns the same loss to the election of either candidate.

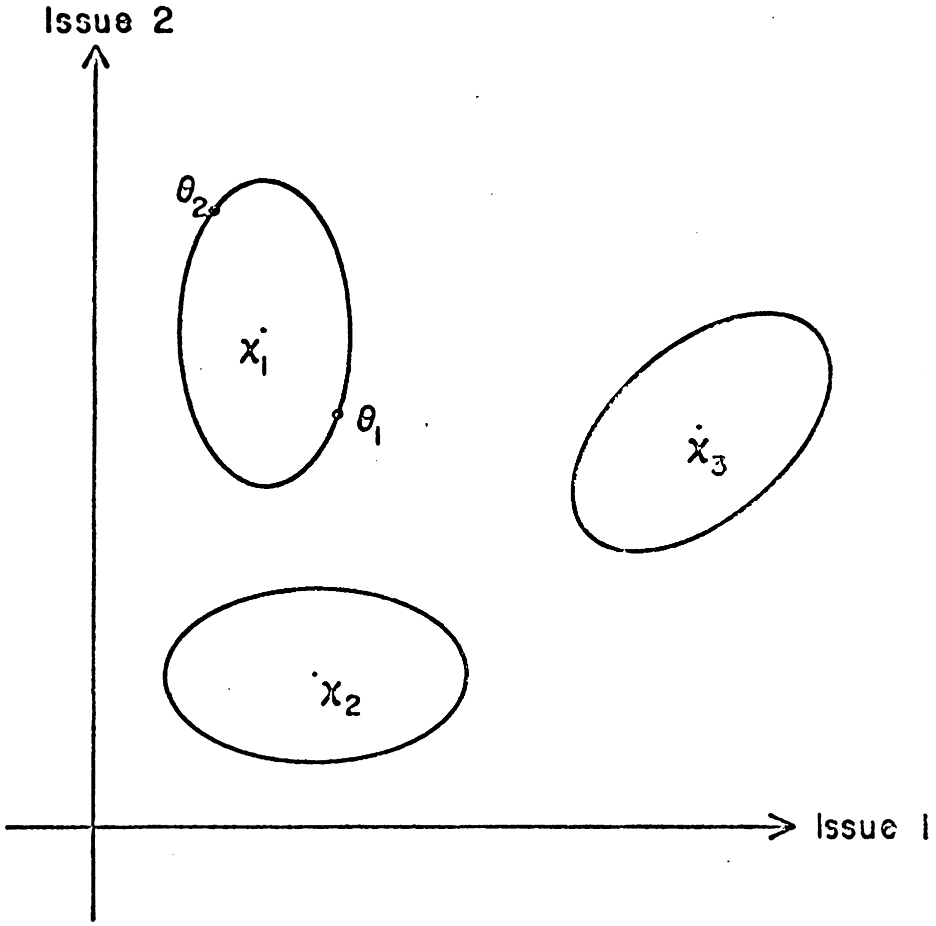



Figure 3.2

Three Voters' Preferred Points and Indifference Ellipses in a Two-Dimensional Issue Space with Two Candidates are known to be $\theta_1$ and $\theta_2$. The more salient of the two issues for $V_1$ is Issue 1, while Issue 2 is the more salient of the two for $V_2$. The third voter is more concerned with some linear combination (interaction) of the issues than he is with either one individually. The first voter is indifferent between $C_1$ and $C_2$, so he will choose between them by a predetermined random process, say by flipping a fair coin. Both of the remaining voters prefer $C_1$ to $C_2$, so, under the majority rule decision function, if all citizens vote, the first candidate will win the election.[6]

The second important observation regarding voters' utility functions concerns the possibility of citizen homogeneity in that respect. In particular, if voters are uniform with respect to the relative salience that they assign to the various election issues, then the same symmetric positive definite loss matrix B enters the utility function of each voter. Thus the utility experienced by the $i^{th}$ voter if the $j^{th}$ candidate is elected is

---

[6]Consider $V_3$, for example, whose preferred point is located at $x_3$. Imagine the class of all ellipsoids concentric to $x_3$ and to which the ellipsoid depicted in Figure 3.1 belongs. The voter is indifferent between two candidate strategies on any specific ellipsoid but he will choose between strategies on two different ellipsoids by asking which ellipsoid is contained within the other. Since, in this example, the ellipsoid concentric to $x_3$ and passing through $\theta_1$ is contained within the one concentric to $x_3$ and passing through $\theta_2$, the third voter would experience greater loss if the candidate whose platform is at $\theta_2$ is elected. He will, therefore, vote for the first candidate. By a similar argument, we can see that the second citizen will also vote for the first candidate.

(3.4) $$U_i(\theta_j) = h_i[(x_i - \theta_j)'B(x_i - \theta_j)].$$

Now it is possible to operate on $E^n$ with a linear transformation so that all of the surfaces of constant utility determined by the matrix B, instead of being ellipsoids in general, are n-dimensional <u>spheres</u>. This situation is illustrated in the two-dimensional issue space in Figure 3.3, where ellipses define the voters' contours of indifference <u>before</u> the linear transformation on the points in the issue space and circles define their contours of indifference <u>after</u> the transformation on those points. The purpose of such a transformation is to simplify the utility functions, since the matrix which determines indifference surfaces in the transformed space is the $n \times n$ identity matrix I.[7] Hence, in the transformed space the utility to the $i^{th}$ voter if the $j^{th}$ candidate is elected is

(3.5) $$U_i(\theta_j) = f_i[(x_i - \theta_j)'(x_i - \theta_j)],$$

where $f_i[\cdot]$, a monotone decreasing function of $(x_i - \theta_j)'(x_i - \theta_j)$, is the function composition of $h_i$ and the linear transformation on $E^n$. Since $(x_i - \theta_j)'(x_i - \theta_j)$ is simply the <u>Euclidean distance</u> between the vectors $x_i$ and $\theta_j$, equation (3.5) indicates that in the transformed space the voter's utility is nothing more than a monotone decreas-

---

[7] It is important to know that if voters have heterogeneous loss matrices, then it will be impossible to find a transformation which will simplify the issue space in this manner.

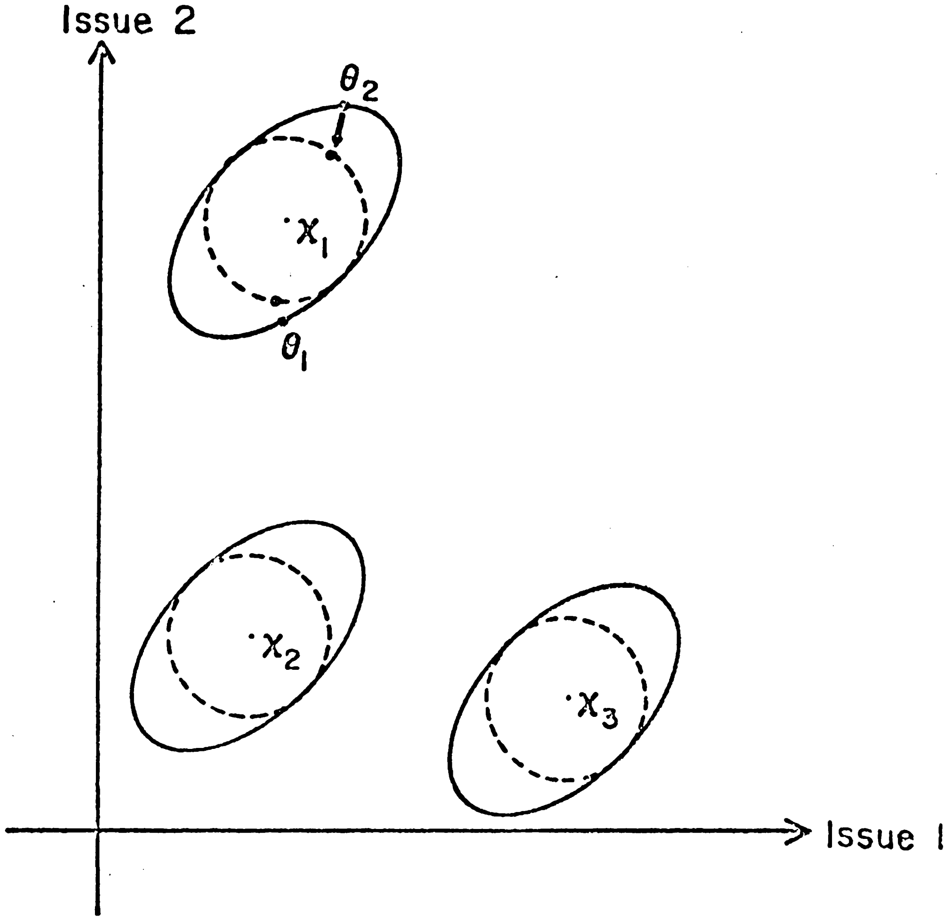



Figure 3.3

Three Voters' Preferred Points and Indifference Ellipses in a Two-Dimensional Issue Space and Their Indifference Circles in a Transformed Space ing function of the distance between his most preferred point $x_i$ and the candidate's platform vector $\theta_j$. In short, if voters share a common loss matrix, then "closeness" or distance will be measured in the Euclidean sense, but if different voters have different loss matrices, then distance will be measured with respect to each particular voter's norm $||\cdot||_{A_i}$. In the latter case, there will be as many metersticks as there are voters in society.

Careful attention should be paid to the fact that the three relations on $E^n$ in Definition 3.1 are specified with respect to a particular voter $(x_i, A_i)$. It is also noteworthy that for each $i$ in some index set $\Lambda$, $I_i$ is an equivalence relation. It would be desirable if, given all of the individual preference orderings $R_i$, $i \in \Lambda$, we could construct a social preference ordering $R$ in a manner that would generally be interpreted to be both rational and consistent with majority rule. An ordering $R$ can be defined by first constructing a distribution or, more formally, a probability measure Pr** on the voter space $E^n \times M_n$, where $E^n$ represents the space of all potential most preferred points and $M_n$ is the space of all $n \times n$ positive definite matrices. Since, for each $i \in \Lambda$, $R_i$ is a complete transitive ordering on $E^n$, if $(X, A)$ represents a voter selected at random from the support of Pr**, we have the following definition of dominance under majority rule.[8]

<u>Definition 3.2</u>: If $y, z \varepsilon E^n$, society prefers $y$ at least as much as $z$, and we abbreviate yRz, if and only if $Pr^{**}(||X - y||_A \leq ||X - z||_A) \geq 1/2$. If society prefers $y$ at least as much as $z$ we say $y$ <u>dominates</u> $z$. If for all $z \varepsilon E^n$, $y$ dominates $z$, then $y$ is said to be a <u>dominant point</u> for the distribution Pr** on $E^n \times M_n$.

By way of example, in a democratic society with a finite number $N$ of citizens the relative frequency assignment of probability is usually intuitively acceptable. That means that the probability of a point $(x_i, A_i) \varepsilon E^n \times M_n$ is simply the number of voters who have preferred point $(x_i, A_i)$ divided by $N$. Then yRz will hold if half or more of the voters are at least as "close" to $y$ as they are to $z$. It is important to note that in this model citizens vote deterministically provided that the normed distances from their preferred points to the candidates' platform vectors are not equal; i.e., if the candidates' platforms are $\theta_1$ and $\theta_2$ and $||x_1 - \theta_1||_{A_i} < ||x_i - \theta_2||_{A_i}$, then $V_i$ will vote for the first candidate with probability 1. If $||x_i - \theta_1||_{A_i} = ||x_i - \theta_2||_{A_i}$, then the $i^{th}$ voter will choose between the two candidates randomly, say by flipping a fair coin. This

[8]The support of a probability measure is the set of all points which are assigned positive density by the measure.

deterministic interpretation is in contrast to the game theoretic approach in which a specific voter can, for example, vote for the first candidate with probability 0.2 and vote for the second with probability 0.8[9] Unfortunately, as is implied by Arrow's Theorem, it is frequently true that "well-behaved" individual preference orderings do not generate a "well-behaved" social preference, so even if all of the voters in a democracy have complete transitive preference orderings, under certain sets of reasonable assumptions, there is no guarantee that $R$ will be transitive.[10]

We conclude the development of preliminary concepts by proving the following useful lemma.

Lemma 3.1: If $(x_i, A_i)$ represents the $i^{th}$ voter and $\theta_1, \theta_2, z \varepsilon E^n$, then the following are equivalent:

1) $\theta_1 P_i \theta_2$
2) $||x_i - \theta_1||_{A_i} < ||x_i - \theta_2||_{A_i}$
3) $(\theta_1-\theta_2)' A_i (x_i - z) > 1/2(||\theta_1 - z||^2_{A_i} - ||\theta_2 - z||^2_{A_i})$

Proof: That 1 and 2 are equivalent is an obvious consequent of Definition 3.1. To show that 2 and 3 are equiva-

[9] For example, see Melvin Hinich, John Ledyard, and Peter Ordeshook, "A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games," Journal of Politics, Vol. 35, (Feb. 1973), pp. 154-193.

[10] For example, see R. W. Hoyer and Lawrence S. Mayer, "On Social Preference Orderings Under Majority Rule," Econometrica, Vol. 43, (July 1975), pp. 803-806.

lent, first note that

$$||x_i-\theta_1||_{A_i}<||x_i-\theta_2||_{A_i} \Longleftrightarrow ||x_i-\theta_1||^2_{A_i}<||x_i-\theta_2||^2_{A_i} .$$

$$\Longleftrightarrow (x_i-\theta_1)'A_i(x_i-\theta_1)<(x_i-\theta_2)'A_i(x_i-\theta_2).$$

By subtracting and adding $z$ from each vector we get

$$(x_i-z+z-\theta_1)'A_i(x_i-z+z-\theta_1) < (x_i-z+z-\theta_2)'A_i(x_i-z+z-\theta_2).$$

Hence

$$||x_i-z||^2_{A_i} + ||z-\theta_1||^2_{A_i} + 2(z-\theta_1)'A_i(x_i-z)$$

$$<||x_i-z||^2_{A_i} + ||z-\theta_2||^2_{A_i} + 2(z-\theta_2)'A_i(x_i-z).$$

But this inequality is equivalent to

$$||z-\theta_1||^2_{A_i} - ||z-\theta_2||^2_{A_i} < 2(\theta_1-\theta_2)'A_i(x_i-z),$$

and the desired result follows at once.

Q. E. D.

## Section 3.2 A Two-Candidate Election with Common Indifference Surfaces

In this section we assume that each citizen in a particular social structure has a unique most preferred position and that all citizens have indifference surfaces of the same shape, i.e., they have a common loss matrix B. In addition, we suppose that there are two candidates running for office and there are no voter abstentions. Since the probability measure will have all of its mass concentrated at the single matrix $B\varepsilon\mathcal{M}_n$, it can actually be considered to be the distribution of the random variable X which maps each voter in the space $V$ of all voters into his most preferred point in $E^n$. In order to distinguish this probability function from the one in Definition 3.2, it will be denoted by Pr*. Furthermore, we will let $\theta_1$ and $\theta_2$ represent the platforms of the first and second candidates respectively.

<u>Definition 3.3</u>: If all voters have a common loss matrix B, then a point $\theta \varepsilon E^n$ satisfying the condition that for any $z \varepsilon E^n$, $Pr^*(||X - \theta||_B \leq ||X - z||_B) \geq 1/2$ is said to be a <u>dominant point</u>.

The attractiveness of a dominant point to a candidate for elective office is obvious. If the first candidate chooses a dominant point as his political platform, then the second candidate must either lose the election or also choose a dominant strategy, in which case the election will be decided by a random mechanism. Unfortunately, there are many distributions of the points in $E^n$ for which no dominant point exists, and there are still other distributions in which a dominant point exists but is not unique.

Another political strategy of interest in this spatial analysis is one that minimizes the expected loss experienced by the voters if the candidate adopting that platform is elected. Davis and Hinich maintain that this is the optimal strategy for a beneficent dictator.[11] It is optimal in the sense that if the beneficent dictator adopts the point of minimal societal loss, then the political environment has the property that on the average the citizenry will experience the least amount of loss. Of course, there are other rational interpretations of the manifestation of beneficent dictatorships; consequently, one shoud not expect the minimal societal loss vector to be the only socially (as opposed to strategically) desirable point in the spatial model. At any rate, we have

Definition 3.4: A point $\theta \varepsilon E^n$ satisfying the condition that for any $z \varepsilon E^n$, $E[||X - \theta||_B^2] \leq E[||X - z||_B^2]$, where E is the expected value with respect to the distribution of preferred points, is said to be a point of minimal societal loss.

Fortunately (for society) any realistic distribution of preferred points will have a point of minimal societal loss; however, it is possible to construct a spatial distribution of at least a countably infinite number of voters in which the expected value $E[||X - \theta||_B^2]$ does not exist. Unfortunately (for the candidates) the adoption of the point of minimal societal loss as a political strategy will not assure victory in an election.

[11] Davis and Hinich, in Mathematical Applications in Political Science, II.

Definition 3.5: If all voters have a common loss matrix, then the mean preference $\mu$ is the expected value $E[X]$ of the distribution Pr* of preferred points.

Although the mean preference is not generally a dominant strategy, it has received quite a bit of attention, especially under the assumptions that all voters have the same loss matrix B and the distribution of preferred points is symmetric.[12] Its importance even then is somewhat inflated since it dominates only because it is a special case of a special type of median which we define as follows:

Definition 3.6: A point $m \varepsilon E^n$ is a total median if for every $a \varepsilon E^n$, $Pr^*(a'(X-m) \geq 0) \geq 1/2$.

If $a, m \varepsilon E^n$ are fixed, then $H = \{x \varepsilon E^n | a'(x - m) = 0\}$ is a hyperplane in $E^n$ containing m.[13] Consequently, in the discrete case, m will be a total median of the distribution Pr* if and only if each hyperplane containing m has at least half of the voter's preferred points on both closed sides of it.[14] It is called a total median because it is

[12] Otto Davis and Melvin Hinich, "On the Power and Importance of the Mean Preference in a Mathematical Model of Democratic Choice," Public Choice, Vol. 5 (Fall 1968), pp. 59-72.

[13] A hyperplane in $E^n$ is an n-1 dimensional space embedded in $E^n$. For example, in $E^2$ (two-space) a hyperplane is a line; in $E^3$ (three-space) a hyperplane is a plane.

[14] By a closed side of a hyperplane H we mean the union of the points on H itself with those that are strictly on one side of it.

independent of the particular basis used to "coordinatize" the space of preferred points.[15] On the other hand, if we specify a basis for the space of social states and let $m^*$ be the vector of medians of the marginal distributions, then we have a basis-dependent median. If the basis is changed and the new vector $m^{**}$ of medians of the marginal distributions is computed, it will probably be the case that $m^* \neq m^{**}$. We call such basis-dependent vectors partial medians. Under the assumption that there is a specific basis used to "coordinatize" the space of preferred points, Wendell and Thorson call $m^*$ a multidimensional median.[16] Formally, we have the following definition.

Definition 3.7: A point $m^* \varepsilon E^n$ is a partial median if there exists a basis $A = \{a_1, a_2, \ldots, a_n\} \subset E^n$ for the vector space $E^n$ such that for each $a \varepsilon A$, $Pr^*(a'(X - m^*) \geq 0) \geq 1/2$.

The following theorems and examples will show how the five "optimal" strategies defined in this section are related.

Theorem 3.1 (Davis and Hinich): If all voters have the same loss matrix $B$, then the mean preference $\mu$ is the

---

[15] By an issue basis we mean the basis of a vector space which can be used to define the issues.

[16] R. E. Wendell and S. J. Thorson, "Some Generalizations of Social Decisions under Majority Rule," paper presented at the Annual Meeting of the Midwest Political Science Association, 1973.

point of minimal societal loss.[17]

Proof: Recall that the point of minimal societal loss $\theta$ is the point that minimizes expected loss. But

$$E[||x-\theta||_B^2] = E[(x-\theta)'B(x-\theta)].$$

$$= E[(x-\mu+\mu-\theta)'B(x-\mu+\mu-\theta)].$$

$$= E[(x-\mu)'B(x-\mu)] + E[(x-\mu)'B(\mu-\theta)]$$
$$+ E[(\mu-\theta)'B(x-\mu)] + E[(\mu-\theta)'B(\mu-\theta)].$$

$$= E[||x-\mu||_B^2] + E[||\mu-\theta||_B^2].$$

Since $||\mu-\theta||_B^2 \geq 0$ for all $\theta \varepsilon E^n$, we minimize the right hand side of the equality by choosing $\theta = \mu$.

Q. E. D.

If for no other reason, the mean preference is valuable because it provides us with a convenient and constructive method for computing the point of minimal societal loss. Furthermore, if, as Davis and Hinich claim, the beneficent dictator should choose the point of minimal societal loss as his political strategy, then the function $\phi: E^n \to E^1$ defined by

$$\phi(\gamma) = (\gamma - \mu)'B(\gamma - \mu)$$

is a measure of the "nonbeneficence" of the candidate whose

[17] Several results attributed to other theorists are proven in this dissertation, but such proofs are included only when existing proofs are either inaccurate or inelegant.

platform is represented by $\gamma$. If the first and second candidates have strategies represented by $\gamma_1$ and $\gamma_2$ respectively, and $\phi(\gamma_1) < \phi(\gamma_2)$, then the first office seeker may not win the election, but at least he is, in some sense, more positively disposed toward the "average citizen" than is the second candidate.

Theorem 3.2 (Davis, DeGroot, and Hinich): If all voters have a common loss matrix B, then a point $\theta \varepsilon E^n$ is dominant if and only if $\theta$ is a total median.

Proof: If $m$ is a total median, then for every $a \varepsilon E^n$, $Pr^*(a'(X - m) \geq 0) \geq 1/2$. Let $a = B'(m - \theta_2)$, where $\theta_2$ represents any strategy the second candidate would like to assume. Then $Pr^*(a'(X-m) \geq 0) = Pr^*((m-\theta_2)'B(X-m) \geq 0) \geq 1/2$. But $Pr^*((m - \theta_2)'B(X - m) \geq 1/2(-||\theta_2 - m||_B^2))$ $\geq Pr^*((m - \theta_2)'B(X - m) \geq 0) \geq 1/2$. By Definition 3.2 and Lemma 3.1, $m$ is dominant. To show that the condition is necessary, suppose $\theta$ and let $a \varepsilon E^n$. By Lemma 3.1, for every $z, \theta, \theta_2 \varepsilon E^n$, $Pr^*((\theta-\theta_2)'B(X-z) \geq 1/2(||\theta-z||_B^2 - ||\theta_2-z||_B^2)) \geq 1/2$. Therefore the statement must hold for $z = \theta_2 = b$, where $b = (1/p)a$ for some positive integer $p$. Thus, $Pr^*(b'B(X - \theta + b) \geq 1/2||b||_B^2) \geq 1/2$, and $Pr^*(b'(X - \theta + b'Bb \geq 1/2||b||_B^2) \geq 1/2$. But $b'Bb = ||b||_B^2$. Consequently $Pr^*(a'(X - \theta) \geq (2p)^{-1} ||b||_B^2) \geq 1/2$. The result follows by taking the limit as $p \to \infty$.

Q. E. D.

In view of Theorem 3.2, it is apparent that, within the framework of this spatial model, and given that all citizens have the same loss matrix, the total median is a very desirable political platform. Nevertheless, there are many distributions of preferred points for which a total median, and thus a dominant strategy, does not exist.

Example 3.1: Suppose that fifteen citizens have their preferred points distributed on the vertices of an equilateral triangle in $E^2$, with four voters at one vertex, five at another, and six at the third (see Figure 3.4). In addition, assume that the ellipsoids of constant loss are all circles, so that distances are measured in the usual Euclidean sense. Then for this distribution,

i) there is no dominant point;

ii) there is an infinite set of partial medians.

To show there is no dominant point, consider the fact that in $E^2$ hyperplanes are lines. It is obvious from Figure 3.4 that there is no point $m$ such that every line through $m$ will have at least half the voters on both closed sides of it. Consequently, there is no total median for this distribution.

The candidate for office in a space of voter preferences for which no dominant point exists must still choose a political strategy, and choosing a reasonable platform in the absence of a "best" single point may be a complex task.

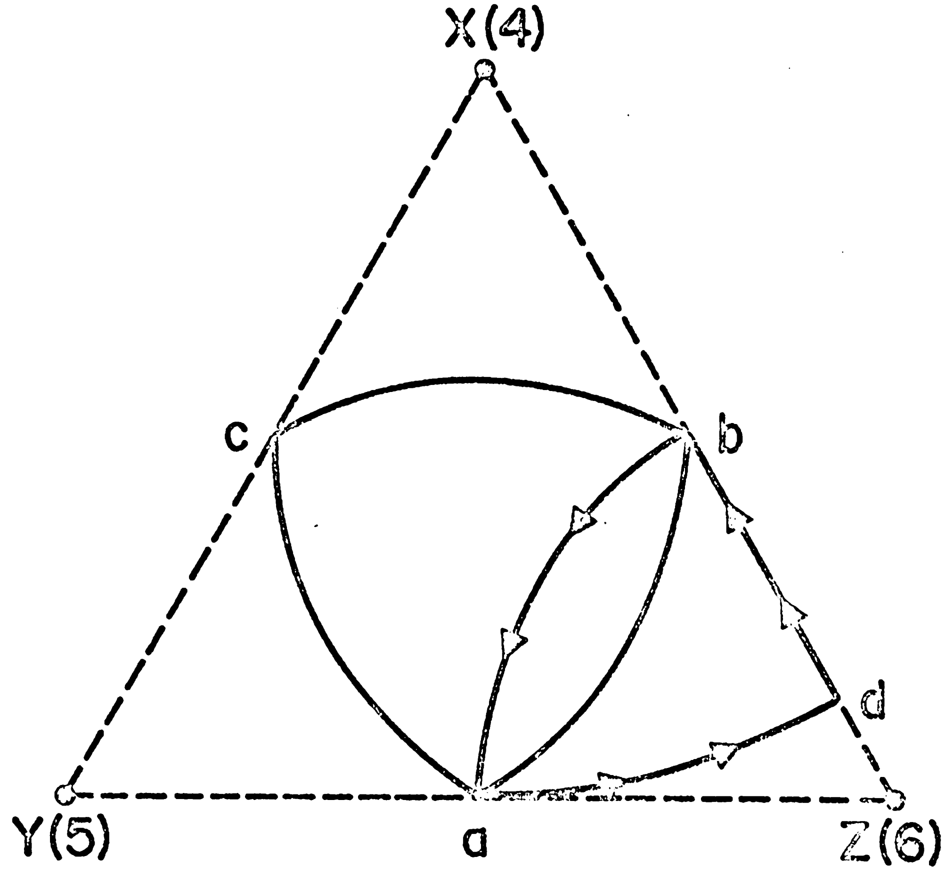



Figure 3.4

A Distribution in Which No Point Is Dominant, the Set of Partial Medians Is Infinite, and There Is a "Large" Intransitive Loop in the Condorcet Set

For example, consider some of the options available in the distribution depicted in Figure 3.4. Regardless of whether a dominant strategy exists, one collection of reasonable strategies is the set of partial medians. A partial median $m^*$ has the positive characteristic that there exists an issue basis such that with respect to each individual issue in that particular basis, $m^*$ is dominant. Notice, by comparing Definitions 3.6 and 3.7, that if $m$ is a total median, it is also a partial median. Now suppose that $m^*$ is a partial median and a dominant point $m$ exists. Then there must be a basis $A$ for which the components of $m^*$ are the medians of the marginal distributions. If $m$ and $m^*$ are distinct points in $E^n$, then they must differ in at least one component. Suppose, for the sake of argument, that they differ in exactly one component. Then the points are identical in every dimension but one, and in that dimension $m^*$ dominates $m$. Since such a situation is impossible (given that $m$ is a dominant point), we must conclude that $m = m^*$. Consequently, if the distribution of social states is known and we know that a dominant point exists, we can find the marginal distributions with respect to the coordinates determined by some basis $A$, and the associated partial median is the dominant point. The object of this development is to devise some constructive procedure for computing a dominant strategy, given that one is known to exist. It is important to recognize that, although the analytic and game theoretic spatial models of majority rule are replete with "existence" theorems, there are very few construction procedures. In Figure 3.4 the closed curve connecting the points a, b, and c is the set of partial medians.

Other platforms that are desirable in the absence of a total median are those in one of the classifications of Condorcet sets.[18] In particular, the Condorcet set C for a distribution of voter preferences in $E^n$ is the minimal (ordered by set inclusion) nonempty set of elements c such that if $z \notin C$, $cPz$. In other words, the Condorcet set is the smallest set of social states with the property that every point in the set is preferred by society to every point not in the set. It has been shown that the partial medians are in the Condorcet set.[19] Furthermore, close analysis of the social preference relation R between points in our example will reveal that, although by restricting his attention to strategies in the Condorcet set the candidate ignores many poor strategies, the Condorcet set is too large to be of much practical use to the candidate. In addition, it is true that any particular point in the Condorcet set, includ-

---

[18] I. J. Good and Lawrence Mayer, "Theorems on Condorcet Sets," (unpublished manuscript, 1972).

[19] R. W. Hoyer and Lawrence Mayer, "Further Results on Condorcet Sets," (unpublished manuscript, 1973).

ing a partial median, frequently has the unfortunate property of being dominated by up to an uncountably infinite number of other points in the preference space. In Figure 3.4 the closed curve connecting a, d, and b is a rather large intransitive loop properly contained in the Condorcet set. As we trace the simple closed curve in the direction of the arrows, say traversing from u to v, then society prefers v to u. Under these conditions, it is conceivable that the optimal strategy may be that of being the last candidate to choose a platform and still have time to inform the electorate accordingly. On the other hand, perhaps the astute office-seeker would simply choose the mean preference $\mu$ and attempt to impress the electorate with his beneficence.

To complete the analysis of the relationship between the five strategies under the assumption of common indifference ellipsoids, we will compare the mean preference $\mu$ with the total median m. First, suppose that $\theta \varepsilon E^n$ and $K \subset E^n$. The set $K\theta = \{2\theta - x | x \varepsilon K\}$ is the reflection of the set K through the point $\theta$.

<u>Definition 3.8</u>: A distribution Pr* on $E^n$ is <u>symmetric</u> about $\theta \varepsilon E^n$ if for each $K \subset E^n$ such that $Pr^*(K) > 0$, $Pr^*(K) = Pr^*(K\theta)$.

<u>Theorem 3.3</u> (Davis, DeGroot, and Hinich): If all voters have a common matrix B and if Pr*, the distribution of the random variable X of preferred points, is symmetric about

$\theta$, then $\theta$ is a total median.

Proof: If the distribution of $X$ is symmetric with respect to $\theta$, then the distribution of $X - \theta$ is symmetric about the vector 0. Thus, for every $a \varepsilon E^n$, $a'(X - \theta)$ is symmetric about the real number 0. By Definition 3.8, $Pr^*(a'(X - \theta) \leq 0) = Pr^*(a'(X - \theta) \geq 0) \geq 1/2$, and we conclude that $\theta$ is a total median.

As a consequence of Theorem 3.3, we see that if the distribution of preferred points is symmetric about the mean preference $\mu$, then $\mu$ is a dominant point. It is not true, however, that symmetry about $\mu$ is a <u>necessary</u> condition for $\mu$ to be a dominant platform.

Example 3.2: Suppose, as is illustrated in Figure 3.5, the distribution of citizens' preferred points in a one-dimensional space can be approximated by the density function $f$ defined by

$$(3.6) \qquad f(x) = \begin{cases} 1/2 & -3/2 \leq x \leq -1/2 \\ 1/4 & 0 \leq x \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

Clearly the distribution is not symmetric about 0. However, in $E^1$ the total medians are simply medians, so $m = \mu = 0$. Consequently, $f$ defines a nonsymmetric distribution in which the mean preference is dominant.

The relationship between points discussed in this sec-

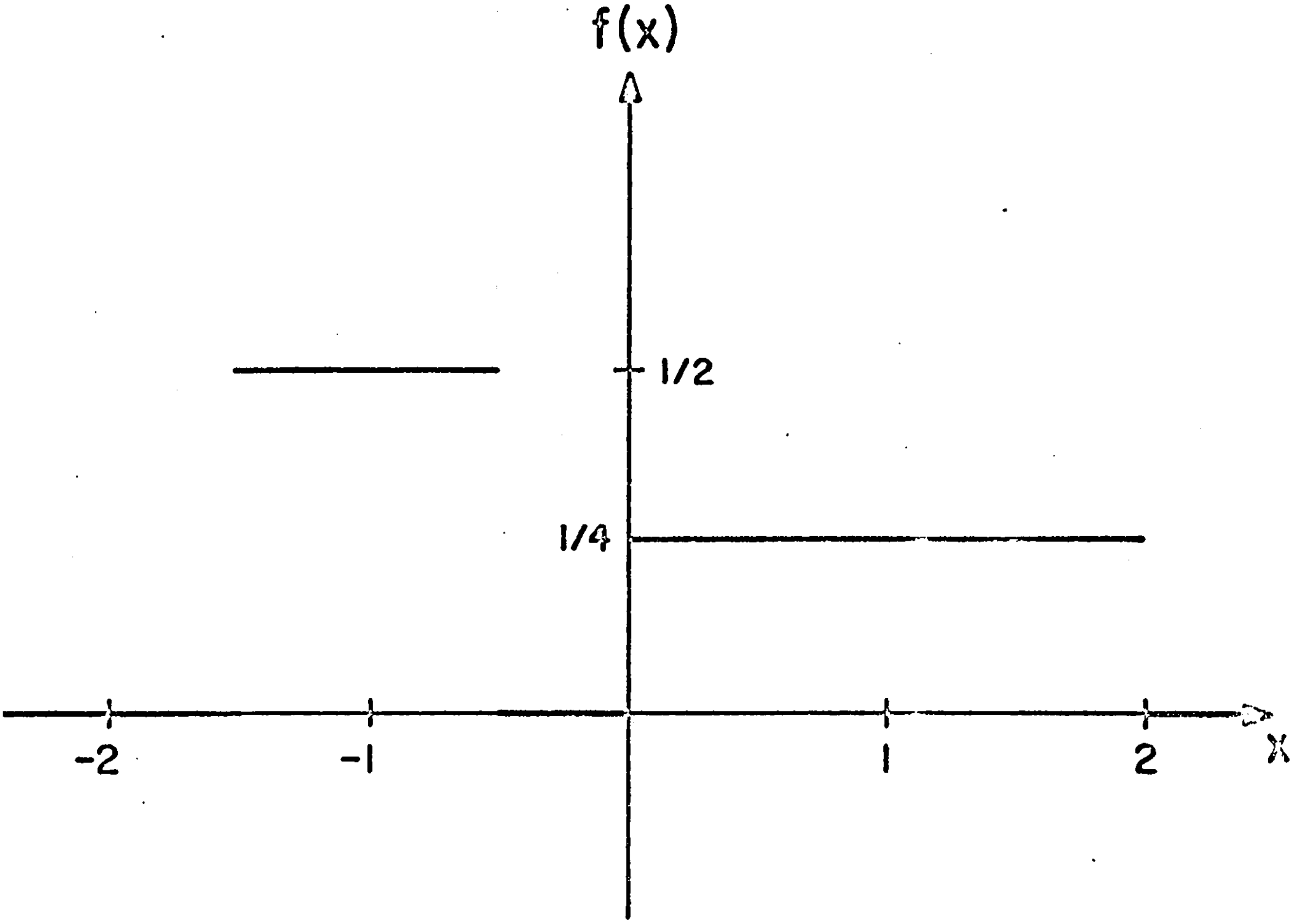



Figure 3.5

A Nonsymmetric Distribution in Which the Mean Preference Is Dominant tion (under the assumption of common loss matrices) is illustrated in Figure 3.6, where the notation $\alpha \xrightarrow[\text{Thm.N}]{Q} \beta$ means that if the point is an $\alpha$ and property Q holds, then the point is also a $\beta$, and the proof can be found in Theorem N. Several of the relationships are not proved as theorems but follow as consequences of the rules of mathematical logic. For example, dominant points and total medians are equivalent concepts; and since every total median is a partial median, it follows that every dominant point is a partial median.

## Section 3.3 A Two-Candidate Election with Individual Indifference Surfaces

At this point, it is useful to characterize the underlying difference between the Davis-Hinich spatial model and our generalized model.

In essence, we would like to move from the Euclidean space $E^n$ consisting of all potential preferred points to the space $E^n \times M_n$ consisting of ordered pairs where the first component is the voter's preferred point in $E^n$ and the second component is the voter's $n \times n$ positive definite loss matrix in $M_n$. Since our statements will be probabilistic ones, it is necessary to define a probability measure Pr** on $E^n \times M_n$, most desirably in such a manner that the measure Pr* used in the previous sections is similar to Pr** for some fixed $B \varepsilon M_n$. This is done by letting $V$ be the space of voters, $B'$ be a $\sigma$-algebra on $V$, and Pr' be a probability

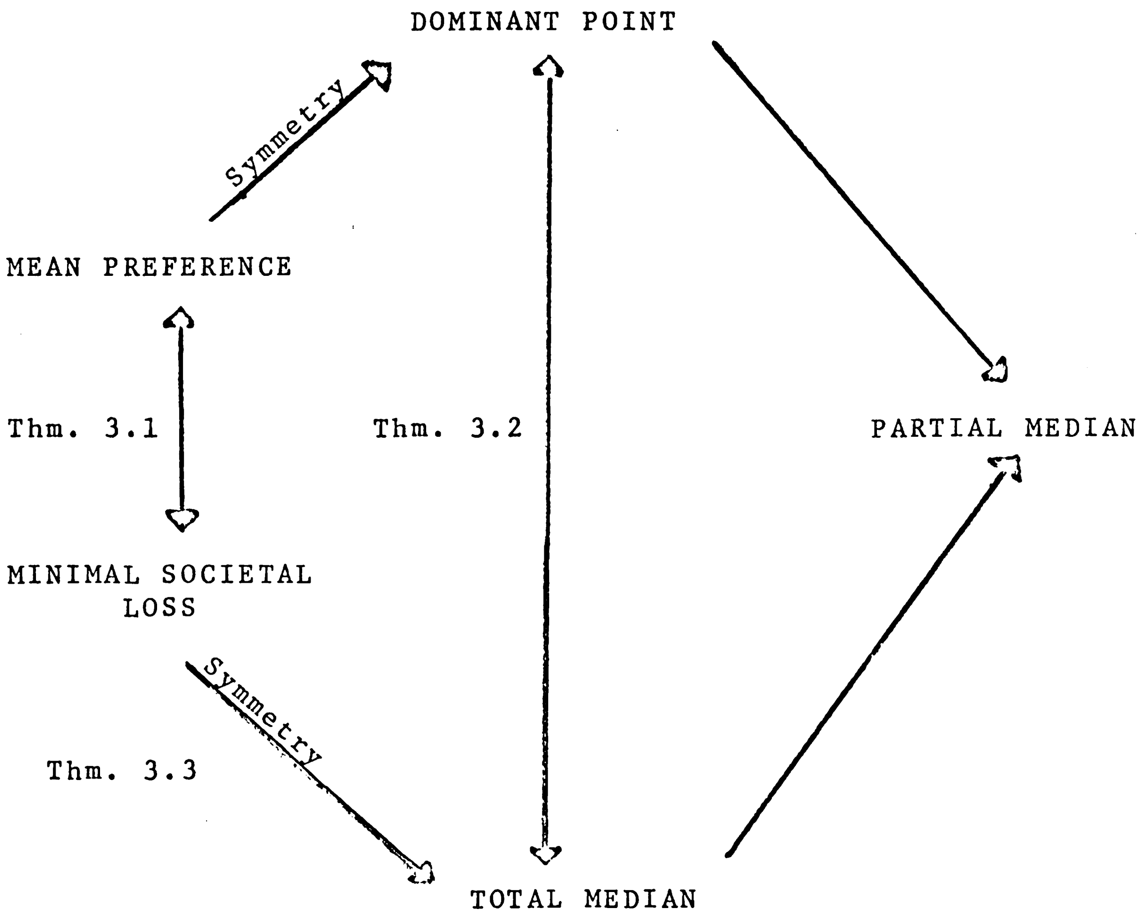



Figure 3.6

The Relationship between Certain "Optimal" Strategies in a Spatial Model Where All Voters Have the Same Loss Matrix measure on $\mathcal{B}'$. Then the random variable of interest in the generalized model is the function (X, A) which maps the $i^{th}$ voter in the space $V$ into the ordered pair $(x_i, A_i) \epsilon E^n \times M_n$. We let $\mathcal{B}$ be the $\sigma$-algebra on the space $E^n \times M_n$ induced by the mapping (X, A), and define the probability measure Pr** on $E^n \times M_n$ by $Pr^{**}(B) = P'((X,A)^{-1}(B))$ for each $B \epsilon \mathcal{B}$.

In other words, if Pr' is a probability measure on the voter space $V$ and the random variable (X, A) tells us precisely what each voter's most preferred point is, then P' induces a probability distribution P** on $E^n \times M_n$. This is the probability function in which we are interested.

Of course, we remain interested in that point $\theta \epsilon E^n$ that minimizes the expected loss of the citizen (X, A) given that the candidate at $\theta$ wins, and believe that such a vector is a reasonable platform for a beneficent dictator.

Definition 3.9: A point $\theta \epsilon E^n$ satisfying the condition that for any $z \epsilon E^n$, $E[||X - \theta||_A^2] \leq E[||X - z||_A^2]$, where E is the expected value with respect to the random variable (X, A), is said to be a point of minimal societal loss.[20]

[20] I. J. Good and Lawrence Mayer, "On Surfaces of Constant Societal Loss in A Model of Social Choice," Journal of Mathematical Sociology, Vol. 2 (August 1972), pp. 209-220.

It is important to note that the difference between this definition and Definition 3.4 is the fact that the relevant random variable (X, A) is a mapping into the product space $E^n \times M_n$, whereas in the previous consideration the random variable was essentially vector-valued in $E^n$.

Definition 3.10: A point $m \varepsilon E^n$ is a generalized total median (GTM) of the distribution Pr** if for every $a \varepsilon E^n$, $Pr^{**}(a'A(X - m) \geq 0) \geq 1/2$.

It may be useful from a conceptual point of view to consider each point $a \varepsilon E^n$ as a map from the space $E^n \times M_n$ into $E^1$ defined by $a(x, A) = a'Ax$. Then for a fixed value of $a \varepsilon E^n$, the event $\{(X,A) | a'A(X - m) \geq 0\} = \{(X,A) | a'AX \geq a'Am\}$ is seen to be a set defined by a relationship between two random variables in a one-dimensional space. It will be convenient throughout the remainder of this section to adopt the notation that $E(AX) = \overline{AX}$ and $E(X) = \overline{X} = \mu$.

Theorem 3.4: In the voter space $E^n \times M_n$, $M = (\overline{A})^{-1}(\overline{AX})$ is the point of minimal societal loss.[21]

Proof: We must prove that $M$ is that value of $\theta$ which minimizes $E[||X - \theta||_A^2]$.

---

[21] In proving this theorem we make use of the following results:

i) $A' = A$.

ii) If A is a random variable in $M_n$, then $E(A) = \overline{A} \varepsilon M_n$.

iii) If $A \varepsilon M_n$, then A is nonsingular.

$$E[||X - \theta||_A^2] = E[(X - \theta)'A(X - \theta)].$$

$$= E[(X - M+M - \theta)'A(X - M+M - \theta)].$$

$$= E[||X - M||_A^2] + E[||M - \theta||_A^2]$$
$$+ E[(X - M)'A(M - \theta)] + E[(M-\theta)'A(X - M)].$$

We will prove that the last summand on the right is 0, and a similar argument will demonstrate that the third term is also 0.

$$E[(M-\theta)'A(X-M)] = E[((\overline{A})^{-1}(\overline{AX}) - \theta)'A(X - (\overline{A})^{-1}(\overline{AX}))].$$

$$= E[((\overline{A})^{-1}(\overline{AX}) - \theta)'(AX - A(\overline{A})^{-1}(\overline{AX}))].$$

$$= ((\overline{A})^{-1}(\overline{AX}) - \theta)'E[AX - A(\overline{A})^{-1}(\overline{AX})],$$

since $(\overline{A})^{-1}(\overline{AX}) - \theta$ is a constant. But $E[AX - A(\overline{A})^{-1}(\overline{AX})] = \overline{AX} - \overline{A}(\overline{A})^{-1}(\overline{AX}) = 0$. Therefore $E[||X-\theta||_A^2] = E[||X - M||_A^2] + E[||M - \theta||_A^2]$. Since $||M - \theta||_A^2 \geq 0$ for all $(\theta, A)\varepsilon E^n \times \mathcal{M}_n$, we minimize the expected value by choosing $\theta = M = (\overline{A})^{-1}(\overline{AX})$, which is the desired result.

Q. E. D.

One striking feature of the expression $(\overline{A})^{-1}(\overline{AX})$ is that it almost appears to reduce to the mean of the distribution of voter's preferred points. In fact, if the random variables $A$ and $X$ are uncorrelated, we have $M = (\overline{A})^{-1}(\overline{AX}) = (\overline{A})^{-1}\overline{A}\mu = \mu$. Furthermore, if $M = \mu$, we have

$M = (\overline{A})^{-1}(\overline{AX}) = \overline{X} = \mu$. Thus $(\overline{A})^{-1}(\overline{AX}) = \mu$. Premultiplying both sides of the equality by $\overline{A}$ yields $\overline{AX} = \overline{A}\mu = \overline{A}\,\overline{X}$. We have proved the following theorem.

<u>Theorem 3.5</u>: In the voter space $E^n \times M_n$, the point of minimal societal loss M and the mean preference $\mu$ coincide if and only if the random variables X and A are uncorrelated.

We have previously asserted that the importance of the mean preference as a rational candidate strategy has been somewhat exaggerated. Even under the assumption that all voters have common indifference ellipsoids, the mean preference will not generally dominate, but, at least in that case, $\mu$ minimizes the expected societal loss. In view of Theorem 3.5, it is apparent that the importance of $\mu$ as an "optimal" strategy has been further diminished. Certainly it is usually unrealistic to assume that all voters are homogenous with respect to the relative importance they assign to the individual campaign issues. To make the additional supposition that the shape of the citizen's indifference ellipsoid is <u>not</u>, on the average, a function of the location of his most preferred point places unfortunate restrictions on the model. To claim that the random variables X and A are uncorrelated means that one should expect, for example, to see voters with "moderate" political positions have loss ellipsoids that are on the average identical to the loss ellipsoids of voters whose most preferred points represent a radical fringe. Our own intuitive feeling is that voters whose preferred points are on the fringe of the distribution of social states frequently have very strong feelings about certain dimensions of the issue space. Consequently, it is unlikely that in general the random variables X and A are uncorrelated. We would not argue that the mean preference is <u>never</u> important; we simply believe that when it appears to be playing a key role as an election strategy, that role can <u>usually</u> be attributed to unrealistic assumptions underlying the model.

The next theorem is the generalized analog of Theorem 3.2, and the proofs of the two theorems are quite similar.

<u>Theorem 3.6</u>: A point $m \varepsilon E^n$ is dominant if and only if m is a generalized total median.

<u>Proof</u>: Suppose m is a GTM. By letting $a = m - \theta_2$ in Definition 3.10, it follows that m is dominant. Conversely, suppose $\theta_1$ is dominant and let $a \varepsilon E^n$. By Lemma 3.1, for every $a, \theta_1, \theta_2 \varepsilon E^n$,

$Pr^{**}((\theta_1 - \theta_2)'A(X - z) \geq 1/2(||\theta_1 - z||_A^2 - ||\theta_2 - z||_A^2)) \geq 1/2.$

In particular, let $z = \theta_2 = \theta_1 - b$, where $b = (1/p)a$ for some positive integer p. Then

$Pr^{**}(b'A(X - \theta_1 + b) \geq 1/2||b||_A^2) \geq 1/2$ or equivalently

$P^{**}(b'A(X - \theta_1) \geq -1/2b'Ab) \geq 1/2$. Therefore

$P^{**}(a'A(X - \theta_1) \geq -(2p)^{-1}a'Aa) \geq 1/2$. The desired result, $Pr^{**}(a'A(X - \theta_1) \geq 0) \geq 1/2$, follows by taking the limit as $p \longrightarrow \infty$.

Q. E. D.

Insofar as the spatial model is concerned, the generalized total median is the ultimate in strategies. Of course, the voter space illustrated in Figure 3.4 will suffice to demonstrate that for certain distributions GTM's may not exist. Further, in the generalized model, the search for "good" platforms in the absence of a GTM is even more complex and inconclusive than was the case when all voters were assumed to share a common quadratic loss function.

Corollary 3.1: If the conditional distribution of preferred points, given any matrix $B$, has a total median $m$ and $m$ is mathematically independent of $B$, then $m$ is dominant.

Despite misgivings about the strategic importance of the mean preference $\mu$, the next theorem gives a sufficient condition for $\mu$ to be a dominant point.

Theorem 3.7: If, in the voter space $E^n \times M_n$, the random variable $Z = A(X - \mu)$ is symmetric about the vector $0$, then $\mu$ is a GTM.

Proof: If $Z$ is symmetric about $0$, then for every $a \varepsilon E^n$, $Pr^{**}(a'Z \geq 0) \geq 1/2$. Hence $Pr^{**}(a'A(X - \mu) \geq 0) \geq 1/2$, so $\mu$ is the generalized total median.

Q. E. D.

It is not surprising that many of the theorems derived from the Davis-Hinich assumptions are no longer true in the generalized setting. It has already been demonstrated in Theorem 3.5 that the mean preference is not necessarily the platform of minimal societal loss. In the next example it will be apparent that the symmetry of the distribution of preferred points must, in fact, be coupled with the homogeneity of weights on the campaign issues in order to <u>guarantee</u> the existence of certain optimal strategies.

<u>Example 3.3</u>: In this example we construct a voter space in which the preferred points are symmetric about a point $x_2$ but $x_2$ is neither a dominant point nor a point of minimal societal loss. In particular, in $E^2 \times M_2$ we consider the three voters $(x_i, A_i)$ for $i = 1, 2, 3$ defined by

| | $A_i$ | $x_i$ | $A_i x_i$ |
|---|---|---|---|
| First Voter | $\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$ | $\begin{bmatrix} -2 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} -10 \\ 6 \end{bmatrix}$ |
| Second Voter | $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ |
| Third Voter | $\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 10 \\ 6 \end{bmatrix}$ |

This distribution is illustrated in Figure 3.7 where a representative indifference ellipse has been drawn for each voter. Of course, the indifference ellipses for the second voter are simply circles concentric about the origin.

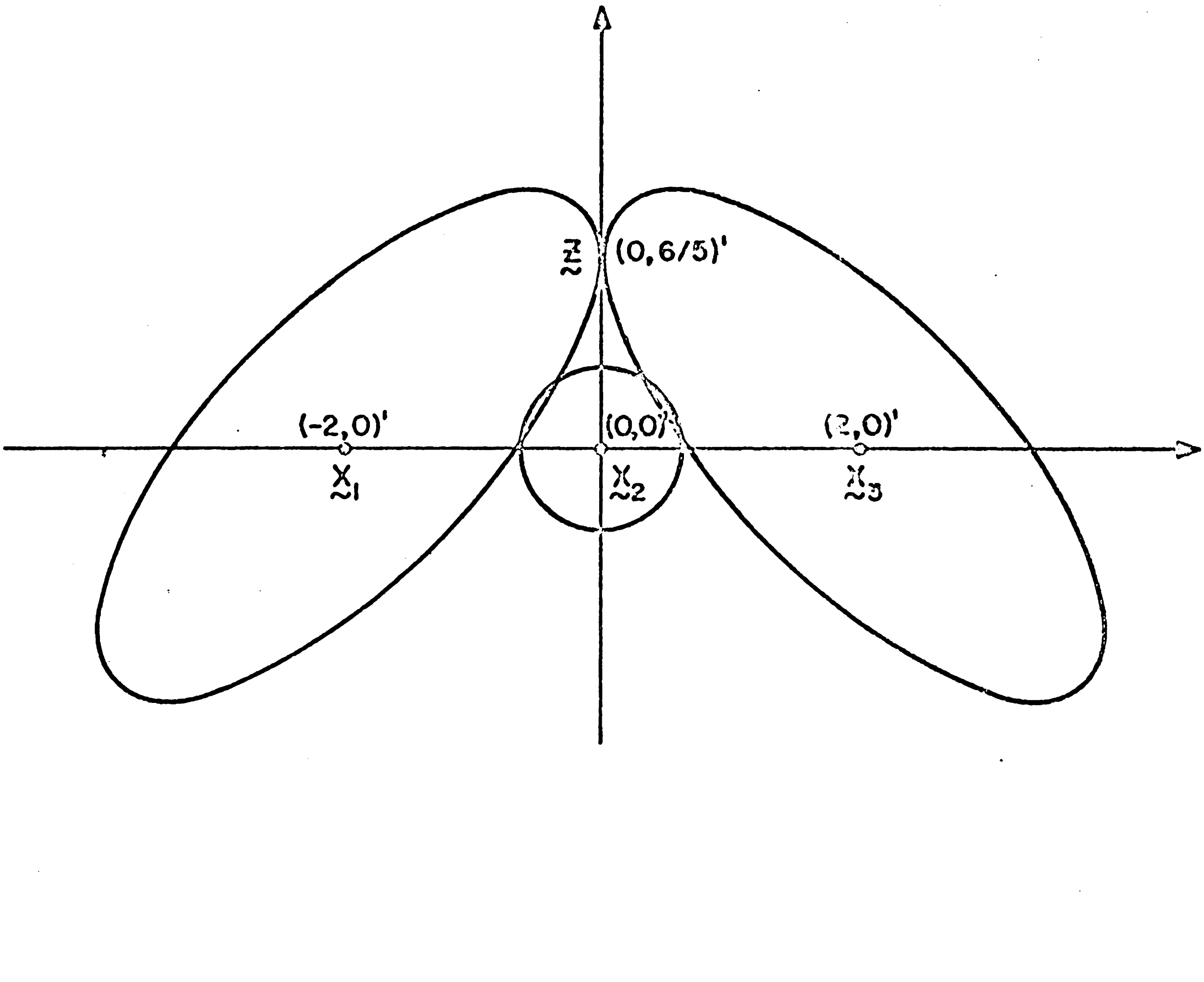



Figure 3.7

A Symmetric Distribution with No Dominant Point and in Which the Mean Preference Is Not the Point of Minimal Societal Loss

Although the preferred points are symmetrically distributed about $x_2$, it can readily be seen that if one candidate is at $x_2$ and the second is at $z$, then the second candidate will get two votes and win the election. Therefore $x_2$ is not a dominant point. Furthermore, a simple computation will yield $M = (\overline{A})^{-1} (\overline{AX}) = [0, 12/11]'$. Hence, the point about which the distribution is symmetric is not the point of minimal societal loss $M$.

At least one area for concern in attempting to fit a mathematical model to a sociopolitical environment and then generalizing the model is whether or not you eventually put yourself out of business. Certainly our principal interest is locating winning political strategies. If, in allowing each voter to define his loss in a completely personal manner, we significantly limit the candidate's ability to locate optimal strategies, then our effort is, for the most part, wasted. In this generalized model the precise opposite is the case. Not only is every total median a generalized total median, but there are distributions for which a GTM exists and a total median does not.

Example 3.4: Consider the voter subspace $E^2 \times M_2$ consisting of five citizens $(x_i, A_i)$, $i = 1, \ldots, 5$ defined as follows:

| | $A_i$ | $x_i$ | $A_i(x_i - \mu)$ |
|---|---|---|---|
| First Voter | $\begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$ | $\begin{bmatrix} -5 \\ -3 \end{bmatrix}$ | $\begin{bmatrix} -7 \\ -5 \end{bmatrix}$ |
| Second Voter | $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ | $\begin{bmatrix} -1 \\ -9 \end{bmatrix}$ | $\begin{bmatrix} -9 \\ -21 \end{bmatrix}$ |
| Third Voter | $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} 2 \\ -2 \end{bmatrix}$ | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ |
| Fourth Voter | $\begin{bmatrix} 7 & 0 \\ 0 & 1 \end{bmatrix}$ | $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ | $\begin{bmatrix} 7 \\ 5 \end{bmatrix}$ |
| Fifth Voter | $\begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$ | $\begin{bmatrix} 11 \\ 1 \end{bmatrix}$ | $\begin{bmatrix} 9 \\ 21 \end{bmatrix}$ |

The distribution is illustrated in Figure 3.8. It is easy to see that there is no point in $E^2$ such that <u>every</u> hyperplane (line) through it has three or more points on both closed sides of it. Consequently, even if all voters had a common quadratic loss function, there is no total median. Nevertheless, the distribution of the variable $Z = A(X - \mu)$ is symmetric about the vector $0$; so by Theorem 3.7 the vector $\mu = [2, -2]'$ is a GTM. It is noteworthy that in this example the point of minimal societal loss is the vector $M = [2, -34/13]'$. The analytic significance, if not the political impact, of this example is obvious. First, if all the voters adopt the same loss matrix, then every GTM reduces to a total median; so it

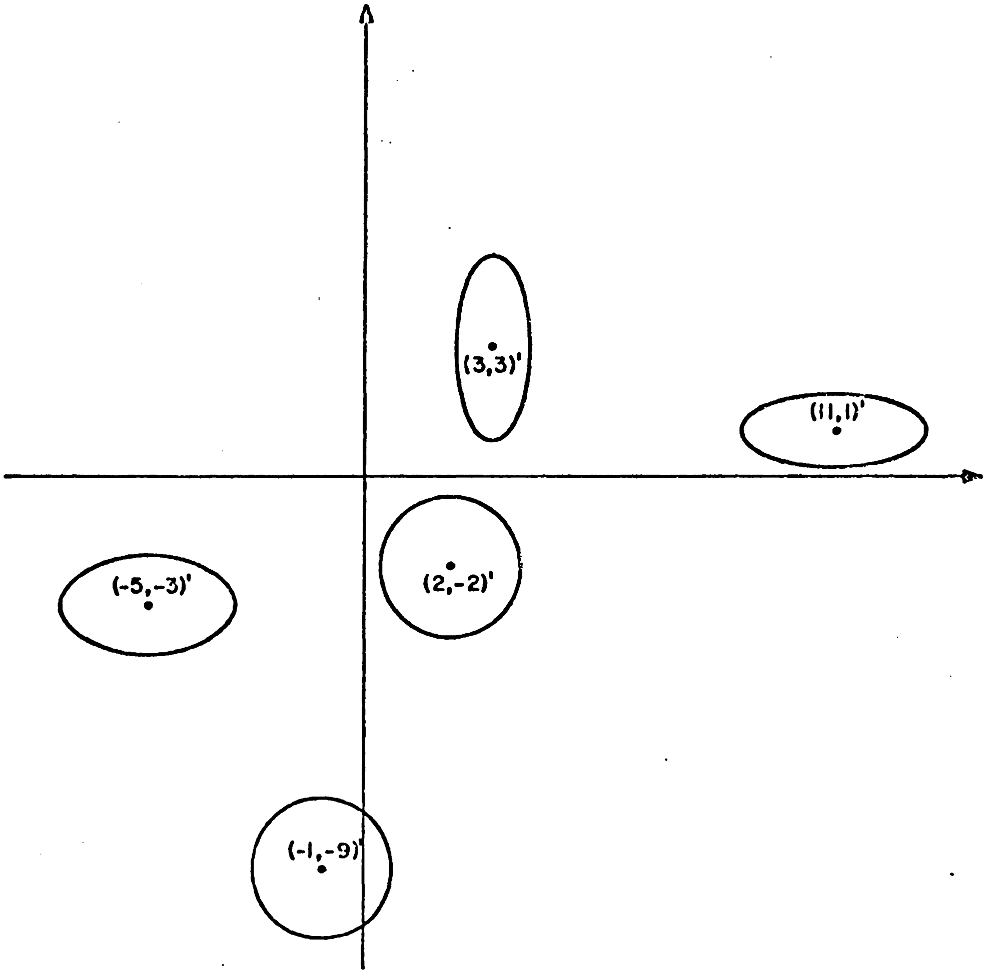



Figure 3.8

A Distribution Which Has a Generalized Total Median but Does Not Have a Total Median follows that every dominant spatial strategy in the Davis-Hinich sense is also a dominant point in the more general setting. Furthermore, there are distributions of voters in the extended model admitting dominant strategies which are GTM's but not Davis-Hinich dominant points. Finally, the point which minimizes the expected loss of the citizenry $M = [2, -34/13]'$ is not the mean preference $\mu = [2, -2]'$. There is no formal reason, other than that the example is contrived, to explain the fact that the GTM is the mean preference.

The relationship between the points discussed in this section is illustrated in Figure 3.9, and the notation is identical to that used in Figure 3.6.

We conclude this chapter with several observations relative to the transitivity of the image of a social welfare function within the context of the generalized model.

First, notice that <u>even if all voters have identical loss matrices B</u>, a combination of symmetry and the existence of a dominant point is not sufficient to guarantee a transitive social preference ordering R under majority rule.

<u>Example 3.5</u>: In Figure 3.10 it is obvious that the distribution of Pr* is symmetric with respect to $\theta$. If the matrix B is defined such that $||\cdot||_B$ is Euclidean distance, then by Definition 3.3 we have $\theta' P x$ and $x R \theta$.

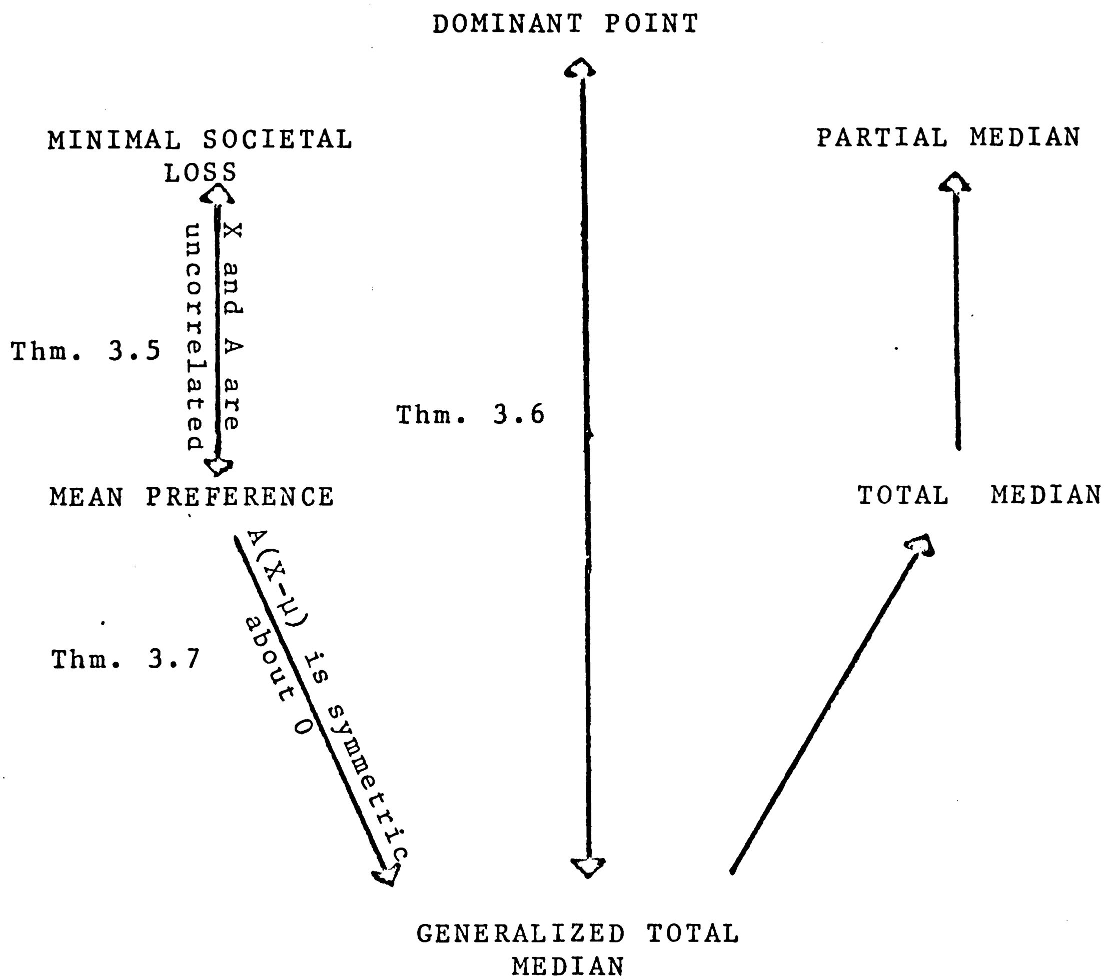



Figure 3.9

The Relationship between Certain "Optimal" Strategies in a Spatial Model Where Each Voter Has His Own Individual Loss Matrix

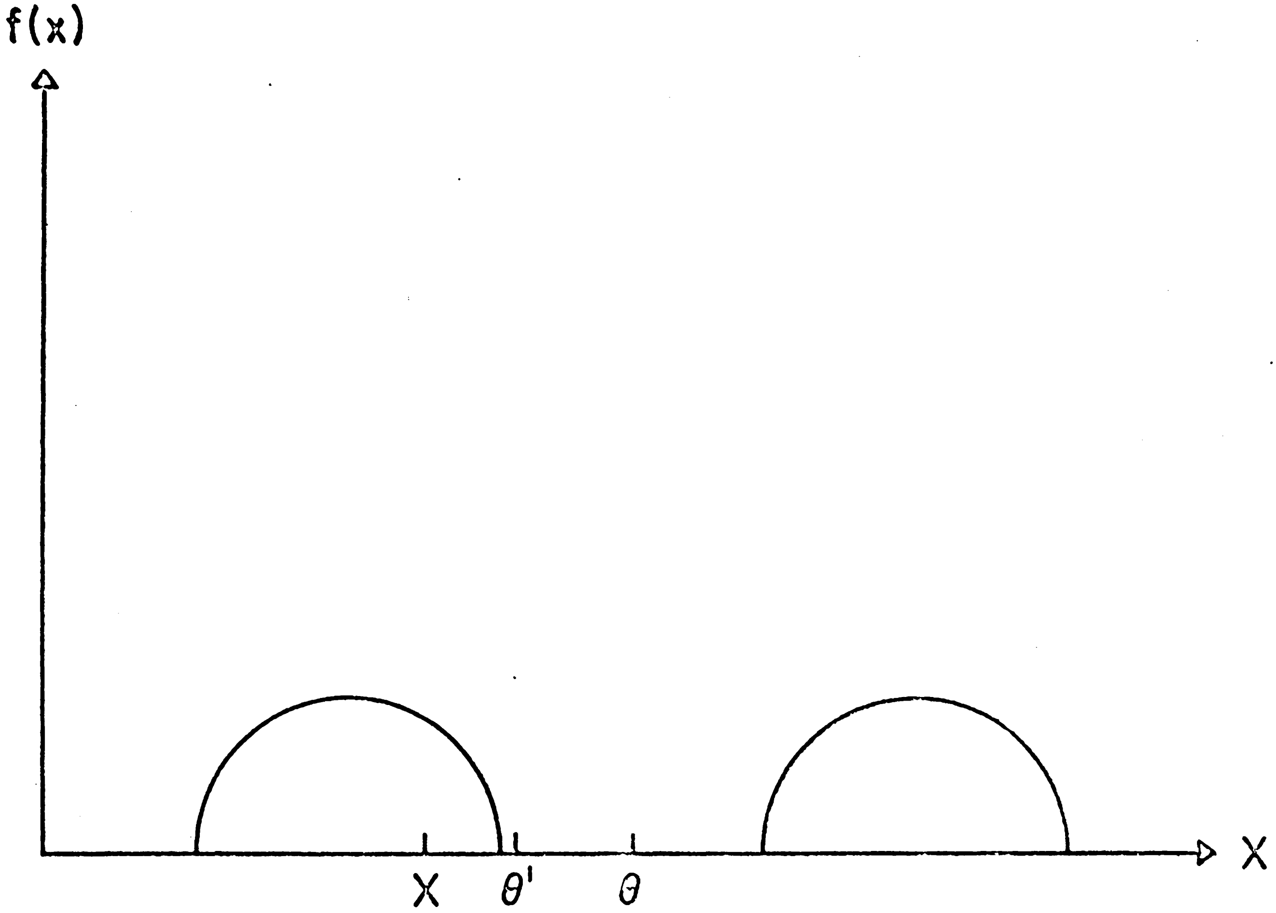



Figure 3.10

An Intransitive Social Preference Ordering Defined on a Symmetric Distribution with a Dominant Point

Furthermore, $\theta$ and $\theta'$ are medians of Pr*, so they are both dominant points and $\theta I \theta'$. In summary, $\theta' R x$, $x R \theta$, and $\theta R \theta'$ all hold. Since society is not indifferent with respect to these pairs, R is not transitive.

It is natural to seek formal restrictions on the voters' preference orderings which insure a transitive social preference ordering R. Davis, DeGroot, and Hinich have shown that if all voters have identical loss matrices B, and if the distribution of preferred points has a dominant strategy and a unique median in all directions, then the social preference ordering is transitive.[22] A unique median in all directions is defined as follows:

<u>Definition 3.11</u>: The distribution Pr* on $E^n$ has a <u>unique median in all directions</u> if there exists a unique $b \varepsilon E^1$, such that for every $a \varepsilon E^n (a \neq 0)$, $Pr^*(a'X \leq b) \geq 1/2$ and $Pr^*(a'X \geq b) \geq 1/2$.

Unfortunately, in the more general setting the existence of a dominant point and a unique median in all directions is not sufficient to guarantee that R is transitive.

<u>Example 3.6</u>: In Figure 3.11, suppose that x, y, and z are each the most preferred point for two voters and $\theta_1$

[22] Otto Davis, Morris DeGroot, and Melvin Hinich, "Social Preference Orderings and Majority Rule," <u>Econometrica</u>, Vol. 40 (January 1972), pp. 147-158.

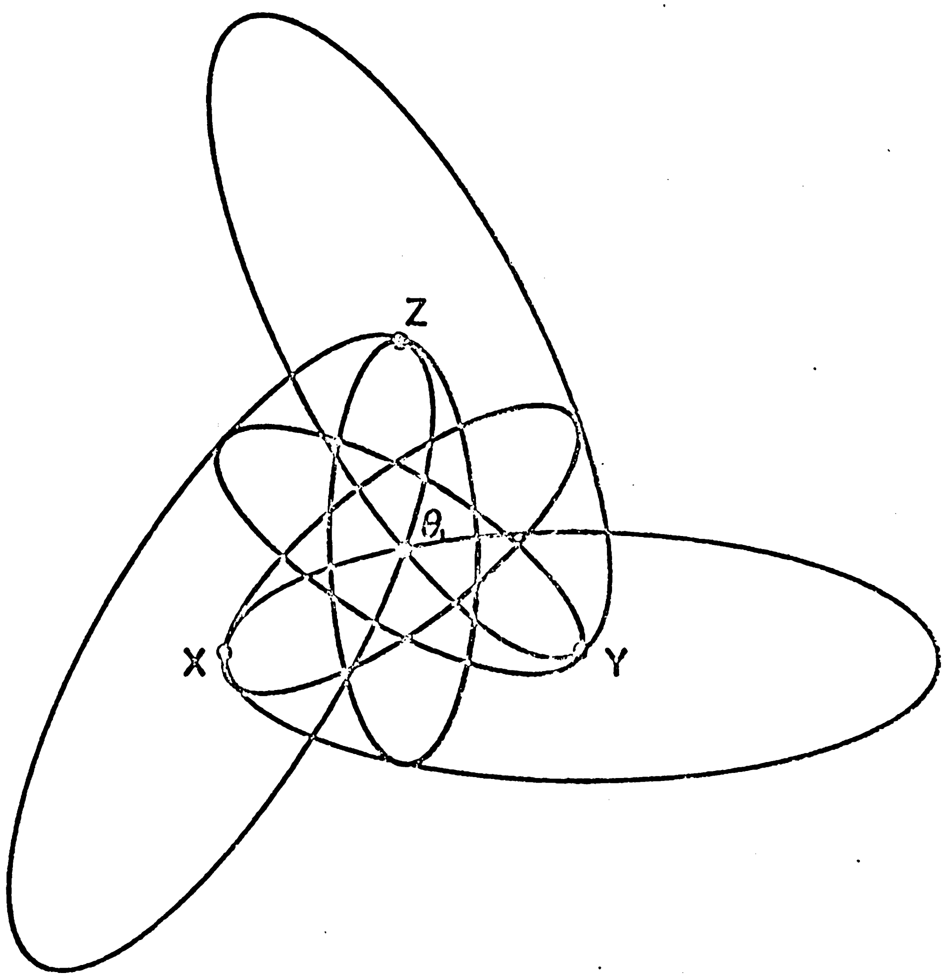



Figure 3.11

An Intransitive Social Preference Ordering Defined on a Distribution with a Dominant Point, a Unique Median in All Directions, and Unique Conditional Medians (on A) in All Directions is the most preferred point for three voters. Let both voters at x have the same matrix with indifference ellipses defined as indicated. The same situation holds for the voters at y and z; however, the three voters at $\theta_1$ have indifference curves as indicated in the display. It is clear that $\theta_1$ is a dominant point since no matter where $\theta_2 \varepsilon E^2$ is located, $\theta_1$ will always receive at least five of the nine votes in a two-candidate race with $\theta_2$. In addition, the distribution has a unique median in all directions since every hyperplane through $\theta_1$ has at least half the voters on both closed sides of it, and any hyperplane not through $\theta_1$ will have less than half of the voters on one closed side of it. Consequently, we have the property that for every $a \varepsilon E^n (a \neq 0)$, there exists a unique $b \varepsilon E^1$ such that $Pr(a'X \leq b) \geq 1/2$ and $Pr(a'X \geq b) \geq 1/2$. It remains to demonstrate that the social preference ordering R is intransitive. In comparing x and y, it is clear that the two voters at y will choose y, the two voters at z will choose y, and one of the three voters at $\theta_1$ will choose y. Since in that binary comparison y is the choice of at least five voters, we have yPx. Similarly, it can be shown that xPz and zPy. Thus P, and consequently R, is intransitive.

# CHAPTER IV

## VOTER ABSTENTION DUE TO ALIENATION AND INTERPERSONAL COMPARISONS

The second limitation of the Davis-Hinich model--that of not providing for citizen abstention from voting induced by alienation--was first treated by Hinich and Ordeshook and has recently received rather broad exposure in Riker and Ordeshook's An Introduction to Positive Political Theory.[1] Unfortunately, the Hinich-Ordeshook analysis is based upon an ambiguous interpretation of an important characteristic of the model, so the problem of how to treat abstentions, even when all voters define their losses in an identical fashion, remains an outstanding one for those interested in spatial models. The purpose of this chapter is to describe explicitly the errors committed by Hinich, Ordeshook, and Riker and make several observations about our own efforts to formulate a model which allows voter abstentions due to alienation.

---

[1]Melvin Hinich and Peter Ordeshook, "Abstentions and Equilibrium in the Electoral Process," Public Choice, Vol. 7 (Fall 1969), pp. 81-106.

Melvin Hinich and Peter Ordeshook, "Plurality Maximization vs. Vote Maximization: A Spatial Analysis with Variable Participation," American Political Science Review, Vol. 64 (September 1970), pp. 772-791.

William Riker and Peter Ordeshook, An Introduction to Positive Political Theory, (Englewood Cliffs, New Jersey: Prentice-Hall, 1973).

As was previously noted, a very desirable characteristic of the D-H model, as well as the generalized model presented in the last chapter, is the fact that it allows one to take individual preference orderings and aggregate them to obtain a social preference ordering without introducing interpersonal comparisons of utility. This point has never been emphasized in the voluminous literature relative to the spatial models of voting theory produced by Aranson, Davis, DeGroot, Hinich, Ledyard, and Ordeshook. Nevertheless, the idea of aggregating individual preferences without admitting interpersonal comparisons is central to the development of virtually all of modern economics (including welfare economics) and, consequently, is fundamental to a substantial portion of both voting theory and the theory of public choice.[2] In fact, if a political participation model incorporates the concept of interpersonal comparisons of utility, then Arrow's paradox does not apply; both the basic definition of his social decision function and one of the value judgments underlying Arrow's Theorem (the independence of irrelevant alternatives axiom) are designed to circumvent analysis based upon such comparisons.

---

[2] We believe that individuals do, in fact, make interpersonal comparisons of utility, and we are willing to argue that the study of such comparisons constitutes a very important research area. Nevertheless, if it is possible to analyze a sociopolotical structure without making interpersonal comparisons of utility, one is almost invariably on more solid terrain for not having made them.

Quirk and Saposnik remark that

> Condition 3 [the independence of irrelevant alternatives] excludes social rankings based upon interpersonal comparisons of utility because intensity of preference is not taken to be relevant to the social ranking--it is only the position of the rankings of individuals that counts in the social ranking. Thus conditions 1 to 5 [the conditions of Arrow's Theorem] exclude both universal social rankings based upon interpersonal comparisons of utility and those of a dictatorial variety.[3]

Sen reports that

> individual utilities are not found in natural cardinal units, and the cardinalization follows experimental observations, yielding a set of numbers that are unique but for an increasing linear transformation. Since the utility scale has to be fixed by specifying the utility value of two points on it, implicitly or explicitly, the other alternatives come

---

[3]James Quirk and Rubin Saposnik, Introduction to General Equilibrium Theory and Welfare Economics, (New York: McGraw-Hill Book Co., 1968), p. 111.

into this valuation. In trying to achieve an interpersonal correspondence, for the sake of social aggregation, this has to be done, and then any use of preference intensity violates not only the 'ordering' aspect of the condition, but also its 'irrelevance' aspect.[4]

To illustrate the manner in which utility specification and independence of irrelevant alternatives are related, consider the following example, a variation of one initially proposed by Arrow.[5] Imagine a collective choice circumstance where three voters $V_1$, $V_2$, and $V_3$ are attempting to choose between social states $o_1$, $o_2$, and $o_3$ by specifying their respective utilities for the three social states and determining the social choice by comparing the sums of the utilities over each social state. Suppose that $V_1$'s preference ordering is $o_1 R_1 o_2 R_1 o_3$ and he assigns utilities of 200, 110, and 100 "utils" to $o_1$, $o_2$, and $o_3$ respectively. As was indicated in the quotation attributed to Sen, the utility numbers are unique up to a linear transformation, so it is common practice to normalize utility specifications

[4] Amartya Sen, Collective Choice and Social Welfare, (San Francisco: Holden-Day, Inc., 1970), pp. 90-91. Also see Riker and Ordeshook, An Introduction to Positive Political Theory, pp. 109-115.

[5] Arrow, Social Choice and Individual Values, p. 32.

between 0 and 1. In particular, for $V_1$ the normalized utilities assigned to $o_1$, $o_2$, and $o_3$ are 1.0, 0.1, and 0.0 respectively. Now assume that $V_2$ and $V_3$ have preference orderings $o_2 R_i o_1 R_i o_3$ and the normalized utilities for all voters are those reported below:

| | | Voters $V_1$ | $V_2$ | $V_3$ | Total Utility |
|---|---|---|---|---|---|
| Normalized Utility | $U_i(o_1)$ | 1.0 | 0.6 | 0.6 | $\sum_{i=1}^{3} U_i(o_1) = 2.2$ |
| | $U_i(o_2)$ | 0.1 | 1.0 | 1.0 | $\sum_{i=1}^{3} U_i(o_2) = 2.1$ |
| | $U_i(o_3)$ | 0.0 | 0.0 | 0.0 | $\sum_{i=1}^{3} U_i(o_3) = 0.0$ |

It is obvious that for this configuration of utility assignments, the social preference is $o_1 P o_2 P o_3$. Suppose, however, that $V_2$ and $V_3$ revise their intensity of feeling with respect to $o_3$, a social state which is an irrelevant alternative in terms of the comparison between $o_1$ and $o_2$, claiming that they are now both indifferent between $o_1$ and $o_3$. Consider the revised normalized utilities for the three social states:

| | | Voters | | | Total Utility |
|---|---|---|---|---|---|
| | | $V_1$ | $V_2$ | $V_3$ | |
| Normalized Utility | $U_i(o_1)$ | 1.0 | 0.0 | 0.0 | $\sum_{i=1}^{3} U_i(o_1) = 1.0$ |
| | $U_i(o_2)$ | 0.1 | 1.0 | 1.0 | $\sum_{i=1}^{3} U_i(o_2) = 2.1$ |
| | $U_i(o_3)$ | 0.0 | 0.0 | 0.0 | $\sum_{i=1}^{3} U_i(o_3) = 0.0$ |

The resultant social preference ordering in the revised setting is $o_2P'o_1P'o_3$. The social preference of $o_1$ over $o_2$ has been reversed by virtue of an alteration in the utilities assigned to an irrelevant alternative, $o_3$. It is noteworthy that in <u>both</u> instances the majority rule preference ordering is $o_2P''o_1P''o_3$. In order to avoid problems similar to the one illustrated here, welfare economists and voting theorists have, for the most part, rejected models incorporating interpersonal comparisons of utility.

We reemphasize, then, that one of the very desirable features of both the D-H model and our generalized model is the fact that they embrace the specification and aggregation of individual <u>preference orderings</u> without requiring specification of individual <u>utilities</u>. On the other hand, if we are interested in the absolute utility received by the $i^{th}$ voter with respect to various social states in $E^n$, then it is imperative that he report, not only his most preferred point and his loss matrix (assuming voters are not homogeneous in this respect), but also his monotone function $h_i$ (see equation (3.2)). However, if our analysis is dependent upon knowing citizens' utility functions, we are at least implicitly making interpersonal comparisons of utility.[6]

## Section 4.1 Abstentions from Voting

Although citizens abstain from voting for a wide variety of reasons, formal modelers have classified those reasons into two general categories labeled indifference and alienation. A voter is indifferent between candidates at $\theta_1$ and $\theta_2$ if his preferred point is "equidistant" from the two. On the other hand, a voter is alienated from the candidates if his preferred point is so far from the closest candidate's strategy that he gets very little utility from the election of that candidate. We would like to incorporate this latter type of abstention into the model, so we

[6] This follows from the fact that once all $h_i$'s have been reported, we can choose any $\theta \varepsilon E^n$ and determine every citizen's utility if a candidate whose campaign strategy is $\theta$ is elected. For example, we may discover that the $i^{th}$ voter gets twice as much utility from a particular choice than does the $k^{th}$ voter. One problem with such comparisons is that they could provide voters who are aware of the spatial structure with some incentive to misrepresent their "true" loss functions in order to gain a strategic advantage for the candidate of their choice. The example presented above is an obvious illustration of a situation in which two of the voters have much to gain by a simple misrepresentation of their "true" intensities of feeling with respect to the available social states or candidates.

assume that

i) all voters have the same loss matrix B.

ii) citizens may abstain from voting.

iii) each citizen's probability of voting is an increasing function of the utility he associates with the election of his preferred candidate.

It is at this point that Hinich and Ordeshook create a problem for themselves that plagues their enterprise from its inception. In order to make statements about the citizen's probability of voting, it is imperative that a probability measure of some sort be defined on the space $E^n \times M_n$. This can be done in a rather large number of ways, but until the measure is defined _explicitly_ we have no recourse but to guess the meaning of statements like, "the probability that the $i^{th}$ citizen votes is 0.7." Unfortunately, Hinich and Ordeshook make probability statements that, on the one hand, appear to be frequentist interpretations of probability and, on the other hand, lead us to the conclusion that the individual citizen is making a personal calculation either by using a personally defined or a globally defined calculus. In any case, they do not define their probability measure, and we are prepared to show that _any_ specification of the probability measures in their model leads to a spatial model encumbered by interpersonal comparisons.

## Section 4.2 Different Probability Formulations

Inasmuch as the difficulty with the Hinich-Ordeshook formulation (hereafter denoted the H-O model) lies in the apparently inconsistent utilization of different probability measures, it seems appropriate at this point to discuss two formal approaches to the definition of "probability," a subjective (or personal) definition and a relative frequency definition. Although there are numerous approaches to defining a probability measure, these two are the ones which must be at least intuitively understood in order to recognize the inconsistency in the H-O model.[7]

First, in specifying a _relative frequency definition of probability_ we invoke the Law of Large Numbers. In crude terms, we define an _event_ and replicate an experiment --one of whose possible outcomes is the event in question-- many times. Then the probability of the event is approximated by the proportion (or relative frequency) of those replications which actually eventuate in the specified event. To define the (frequency) probability of the event

---

[7]In order to motivate this discussion, we will utilize an example from the theory of political participation. Nevertheless, the reader is encouraged to contemplate the difference between these two formulations of a probability measure by imagining what the local weatherperson means by, "The probability [or chance] of rain in southern Ohio during the next twenty-four hours is 80 percent."

we take the limit of these approximations as the number of replications increases without bound.

Now for the candidate who is attempting to determine a spatial political platform at which he maximizes the probability that a randomly selected citizen will cast a vote for him (assuming that all citizens vote), there are some obvious shortcomings of such a probability measure. One problem is that the candidate cannot actually <u>know</u> the probability until the election has taken place. Of course, he could employ one of the numerous polling firms to take a random sample from the electorate and <u>estimate</u> the probability, but that would only be an estimate, and it would only be relevant for the time at which the population was polled. One could possibly obtain a better estimate by decreasing the time interval between the polling date and the election, but that would make it difficult for the candidate to alter his campaign strategy if the probability estimate that a randomly chosen citizen would support him on election day were unacceptable to him. Furthermore, public opinion polls are often expensive.

Second, to determine a <u>subjective (or personal) probability distribution</u>, we need a "rational" individual whose subjective judgment will specify the probability measure. In this case, the individual's subjective probability estimate that some well-defined event will occur is the maximum number of dollars, say, that the individual is willing to wager that the event will occur divided by the sum of that number and the number of dollars that he will lose if the event does not eventuate.

For example, if the "rational" individual is a candidate for some political office, then he could compute the probability that the generic voter $V_i$ will support him by placing an imaginary wager on that event. If he would be willing to lose two dollars if $V_i$ does not support him, provided he will win five dollars if $V_i$ does vote for him, then his probability estimate that $V_i$ will vote for him is $5/(5 + 2) = 5/7$. It is noteworthy that the candidate may determine acceptable betting odds after polling the electorate and estimating the relative frequency probability that a randomly chosen citizen will support his candidacy, but even then there is no reason to believe that the two resultant probabilities will be the same. By way of illustration, a liberal Democratic candidate's subjective estimate that he will win an election in an upper middle class voting district may be 0.4, even though a recent poll shows him winning one-half of the vote, simply because he <u>believes</u> that upper middle class individuals enjoy the appearance of being "liberal," but only occasionally vote in a manner consistent with their expressed beliefs.

Still another conceptualization of probability--the objective definition--specifies that every well-defined event has a "true" probability (the true state of nature) which is known only by the Diety. Of course, being merely mortal, we cannot know with certainty what the so-called "true probability" of the event is, so in order to estimate it we usually resort to some empirical process such as the two previously discussed. In any case, it is important to know that, while it is not impossible, it is extremely unlikely that any given individual would have a subjective probability structure over a non-trivial algebra of events which is consistent with a relative frequency probability structure over those events. The fact that these two probability structures are most likely to be different has important consequences for the H-O formulation of abstentions due to alienation in a model of voting under majority rule.

## Section 4.3 Personal Probability Models

Let us suppose that the probability statements in the relevant voting model are made within the framework of a personal calculus for each citizen. If, in choosing between candidates at $\theta_1$ and $\theta_2$, the $i^{th}$ citizen decides to vote, he will vote for the one at $\theta_1$ only if

$U_i(\theta_1) > U_i(\theta_2)$.[8] In other words, if voters do not abstain, they always vote for the candidate whose election will give them the greatest utility. Now voters not only receive utility from the election of their favorite candidates, they also derive some utility from participating in the act of voting itself. We define a "composite" utility for the $i^{th}$ voter by

$$(4.0) \quad \pi_i(\theta_1,\theta_2)(\omega) = \begin{cases} U_i(\theta_1) + \varepsilon_i(\omega) & \text{if } U_i(\theta_1) \geq U_i(\theta_2) \\ \\ U_i(\theta_2) + \varepsilon_i(\omega) & \text{if } U_i(\theta_1) < U_i(\theta_2) \end{cases}$$

where $\theta_1$ and $\theta_2$ are the candidates' strategies, $U_i$ is the $i^{th}$ voter's deterministic utility function, $\varepsilon_i$ is a random variable with expected value $E[\varepsilon_i(\omega)] = 0$, and $\omega$ is in the domain of $\varepsilon_i$. The essential nature of the stochastic term in (4.0) is not specified by Hinich and Ordeshook, so in this section we will analyze what appear to be the most plausible explanations of $\varepsilon_i$, given that each citizen has a personal voting calculus.

First, suppose that there is a "global" random variable

---

[8] Actually the citizen may vote for the candidate at $\theta_1$ even when $U_i(\theta_1) = U_i(\theta_2)$. In this case, when the voter is indifferent between the two candidates, he will presumably resort to a random process to determine the one for whom he will cast his vote. Since this exception does not affect our analysis, we will ignore it.

$\varepsilon$ and $\varepsilon_i(\omega)$ really means $\varepsilon(\omega)$, the image under $\varepsilon$ of a value $\omega$ sampled by the $i^{th}$ citizen from the domain of the random variable. In this formulation the voter is free to determine a realized value $\varepsilon(\omega)$ of $\varepsilon$ (by reacting to weather conditions, the "importance" of the election, political polls, etc.), but he is not allowed to specify the random variable itself. Quite the contrary, the random variable is a function of the make-up of the total electorate. Under this assumption $\pi_i[\cdot, \cdot]$ is structurally different from $U_i[\cdot]$, because, for each i = 1,..., N, it is a mapping from $E^n \times E^n$ into a space $\Pi$ of random variables. Since for each i and for each fixed pair of candidate strategies $(\theta_1, \theta_2) \varepsilon E^n \times E^n$, $\pi_i(\theta_1, \theta_2)$ is a random variable, it is possible to determine probability measures on the range space of the random variables by simply identifying the measures induced by these mappings. In this formulation, voters have different stochastic utility functions $\pi_i(\theta_1, \theta_2)$; however, these functions differ only in the deterministic term $U_i$. Hence, there is only a single relevant probability measure Pr for all citizens.

To complete analysis of this conceptualization of a personal calculus, we assume that the $i^{th}$ citizen will vote if and only if his utility $\pi_i(\theta_1, \theta_2)$ from voting is positive. We have

<u>Definition 4.1</u>: If the candidates' platforms are represented by the vectors $\theta_1$ and $\theta_2$, if $\pi_i(\theta_1, \theta_2)$ is the utility the $i^{th}$ citizen derives from voting for his favorite candidate, and Pr is the probability measure induced by the random variable $\varepsilon$, then the probability that the $i^{th}$ citizen votes is $\Pr[\pi_i(\theta_1, \theta_2)(\omega) > 0]$.

In short, a citizen's utility from voting $\pi_i(\theta_1, \theta_2)$ is the sum of a deterministic term $U_i$ and a stochastic term $\varepsilon$, and he will vote (not abstain) if and only if his composite utility for a "sampled" value $\omega$ of the random variable $\varepsilon$ is positive. Furthermore, if he does vote, it will be for the candidate at $\theta_1$ only if either $U_i(\theta_1) > U_i(\theta_2)$ or else $U_i(\theta_1) = U_i(\theta_2)$ and his predetermined random selection process dictates that he vote for the candidate at $\theta_1$.

Now, with no loss of generality, we assume that the $i^{th}$ citizen prefers the candidate at $\theta_1$ to the one at $\theta_2$. Observe that

$$\begin{aligned}\Pr[\pi_i(\theta_1, \theta_2)(\omega) > 0] &= \Pr[-\varepsilon(\omega) < U_i(\theta_1)].\\ &= \Pr[-\varepsilon(\omega) < f_i(x_i - \theta_1)'(x_i - \theta_1)].\\ &= \Pr[f_i^{-1}(-\varepsilon(\omega))<(x_i - \theta_1)'(x_i - \theta_1)].\\ &= \Pr[f_i^{-1}(-\varepsilon(\omega)) < ||x_i - \theta_1||^2],\end{aligned}$$

where $||x_i - \theta_1||^2 = (x_i - \theta_1)'(x_i - \theta_1)$ is the Euclidean distance from the $i^{th}$ voter's preferred vector to his favorite candidate's platform. Furthermore, for a given citizen $x_i$ and a given strategy vector $\theta_1$, $||x_i - \theta_1||^2$ is a fixed real number. Hence, the probability that the $i^{th}$ citizen abstains from voting can be written as a stochastic function of the distance from his preferred point to the position $\theta_1$ of his favorite candidate. In particular, we have

$$\Pr[\pi_i(\theta_1, \theta_2)(\omega) > 0] = g_i(||x_i - \theta_1||^2) \tag{4.1}$$

where $g_i[\cdot]$ is a monotone decreasing function of the Euclidean distance between the voter's preferred point and the candidate's strategy.

It is obvious from the previous discussion that we stand accused of being obsessed with subscripts. These subscripts, in addition to emphasizing that the relevant probability measures are <u>personal</u> ones, also call attention to our principal criticism of the H-O formulation of voting where abstentions are allowed. Hinich and Ordeshook have previously claimed that

> $\Pr[\pi > 0] = \Pr[\varepsilon > -U(x, \theta)]$. Assuming that the density of $\varepsilon$ is independent of $x$ and $\theta$, this probability is expressed as a function of $\phi[(x - \theta)'(x - \theta)]$, or simply as $g[(x - \theta)'(x - \theta)]$. Hence $g[(x - \theta)'(x - \theta)] = \Pr[\pi > 0]$.[9]

---

[9]Hinich and Ordeshook, "Abstentions and Equilibrium in the Electoral Process," p. 85.

Then they use the function $g$ to define the probability that a <u>randomly selected</u> citizen whose preferred point is $X$ votes for the candidate $\theta_1$ by

$$(4.2) \qquad V(\theta_1, \theta_2) = \int_S f(X) g[(X - \theta_1)'(X - \theta_1)] dX,$$

where $\theta_2$ is the social state favored by the second candidate, $f$ is the distribution of voters' preferred points in $E^n$, and $S = \{X | \phi[(X - \theta_1)'(X - \theta_1)] < \phi[(X - \theta_2)'(X - \theta_2)]\}$. It should be clear from (4.1) that $V$ is not well-defined, since $g$, instead of being one specific function, is actually one of a class of N, possibly distinct, functions. Furthermore, the formulation cannot be repaired by requiring all voters to have identical $g_i$'s. To see this, suppose that the $i^{th}$ voter's utility function is of the form $\pi_i(\theta_1, \theta_2) = U_i(\theta_1) + \varepsilon$. Then for the $i^{th}$ and $k^{th}$ voters with the same preferred point, who both prefer the candidate at $\theta_1$ to the one at $\theta_2$ we have

$$g_i(||x_i - \theta_1||^2) = g_k(||x_k - \theta_1||^2).$$

$$Pr[\pi_i(\theta_1, \theta_2)(\omega) > 0] = Pr[\pi_k(\theta_1, \theta_2)(\omega) > 0].$$

$$Pr[h_i^{-1}(-\varepsilon(\omega)) < ||x_i - \theta_1||^2] = Pr[h_k^{-1}(-\varepsilon(\omega)) < ||x_k - \theta_1||^2].$$

$$h_i^{-1}(-\varepsilon(\omega)) = h_k^{-1}(-\varepsilon(\omega)).$$

But $h_i$ and $h_k$ defined in equation (3.2) are one-to-one mappings, so $h_i = h_k$. Hence the two citizens' utility functions are identical, and by any definition this constitutes an interpersonal comparison of utility. We conclude that to interpret $\varepsilon$ as a global random variable will lead to unfortunate inconsistencies with the assumptions of the original Davis-Hinich model.

On the other hand, instead of assuming that the stochastic term in (4.0) is a global random variable, we can investigate the possibility that the $i^{th}$ citizen has his own individual random variable $\varepsilon_i$. As in the previous formulation, if we allow each citizen to determine his own calculus of voting by choosing his own random error function $\varepsilon_i$, then there could be up to $N$ distinct stochastic utility functions of interest--one for each citizen in the social order. In this case, if $Pr_i$ is the probability measure induced by the random variable $\varepsilon_i$, we would actually have $N$, possibly distinct, measures defined on different probability spaces (since the $\varepsilon_i$ themselves may be defined on different spaces). Under these conditions it is rather difficult to formalize the concept of an interpersonal comparison of utility; however, from either a structural or substantive perspective, if one citizen's probability of voting is of necessity a function of any other citizen's probability of voting, that would certainly constitute an interpersonal comparison. Now we define

Definition 4.2: If the candidates' platforms are represented by the vectors $\theta_1$ and $\theta_2$, if $\pi_i(\theta_1, \theta_2)$ is the utility the $i^{th}$ citizen derives from voting for his favorite candidate, and $Pr_i$ is the probability measure induced by the random variable $\varepsilon_i$, then the probability that the $i^{th}$ citizen votes is $Pr_i[\pi_i(\theta_1, \theta_2)(\omega) > 0]$.

Notice that the only difference between Definitions 4.1 and 4.2 is the fact that there is a single probability measure for all voters in the former, while each voter has his own individual measure in the latter. As was the case in the previous discussion, it is easy to show that for $i = 1, \ldots, N$,

$$Pr_i[\pi_i(\theta_1, \theta_2)(\omega) > 0] = Pr_i[f_i^{-1}(-\varepsilon(\omega)) < ||x_i - \theta_1||^2].$$

It follows that

$$(4.3) \qquad Pr_i[\pi_i(\theta_1, \theta_2)(\omega) > 0] = g_i(||x_i - \theta_1||^2),$$

where $g_i[\cdot]$ is a monotone decreasing function of the Euclidean distance between the voter's preferred point and the candidate's strategy.[10]

Again, there is not a single function $g_i$, but an entire family of them--one for each voter in the electorate. Hence, even under the assumption that each citizen has his

[10] The $g_i$'s in (4.1) and (4.3) are actually different functions, but we will use the same notation for both in order to be consistent with the H-O terminology.

own personal stochastic term, (4.2) is not well-defined. Furthermore, if we require that all voters have the same $g$ (in order to make (4.2) well-defined), then for the $i^{th}$ and $k^{th}$ voters with identical preferred points and the same favorite candidate, say at $\theta_1$, we have

$$g_i(||x_i - \theta_1||^2) = g_k(||x_k - \theta_1||^2).$$

$$Pr_i[\pi_i(\theta_1, \theta_2)(\omega) > 0] = Pr_k[\pi_k(\theta_1, \theta_2)(\omega) > 0].$$

Since the probability that the $i^{th}$ citizen votes is a function of the probability that the $k^{th}$ citizen votes, an obvious interpersonal comparison of utility ensues when we assume that each voter has his own random variable and they all have identical $g_i$'s. Thus, we conclude that if Hinich and Ordeshook intend for (4.0) to determine a citizen's personal calculus of voting, results inconsistent with the original Davis-Hinich assumptions obtain. In particular, they require interpersonal comparisons of utility.

Finally, we note that Hinich and Ordeshook argue that if the candidate wishes to maximize the number of votes he receives, he should choose as his political platform the vector $\theta_1$ that maximizes the objective function $V(\theta_1, \theta_2)$; whereas the second candidate with an equivalent goal should choose a vector $\theta_2$ which maximizes the function $V(\theta_2, \theta_1)$.

If, on the other hand, the first candidate wishes to maximize his plurality over his opponent, he should select the strategy that maximizes the objective function

$$(4.4) \qquad P(\theta_1, \theta_2) = V(\theta_1, \theta_2) - V(\theta_2, \theta_1).$$

Inasmuch as V is not well-defined, neither is P. Hence, if the relevant probability measures in the H-O formulation of voting--with abstentions--are generated by the individual voters' calculi, either interpersonal comparisons of utility will be introduced into the model or else the candidates' objective functions, as expressed in (4.2) and (4.4), will not be well-defined.

Before examining the final interpretation of their voting model, we observe that in certain contexts, when discussing the probability that a citizen will abstain from voting, Hinich and Ordeshook give the impression that the probability measures in their model are personal ones and are determined by the individual citizens' personal calculi of voting. For example, in their 1970 APSR article (Assumption 8) they state that if we confine our attention to the citizens who prefer $\theta_1$ to $\theta_2$, then

> any two citizens vote with equal probability if their preference vectors (x) are equidistant (as measured by the metric $||x - \theta||_A$) from the strategy $\theta_1$. We assume, moreover,

> that citizen i votes with greater probability than citizen j if and only if the distance between citizen i's preference vector and $\theta_1$ is less than the distance between citizen j's preference vector and $\theta_1$.... This assumption does not imply an interpersonal comparison of utility, but simply that if the preference vectors of two citizens are equidistant from their preferred candidate(s), then the benefits and costs of voting, as well as the density of $\varepsilon$, stand in the same relationship to each other for both citizens.[11]

We have shown that the precise opposite is the case and that such a spatial formulation does, in fact, imply an interpersonal comparison of utility.

## Section 4.4 Relative Frequency Probability Model

In this section we will assume that each citizen in a particular social order decides whether or not to abstain from voting by some personal non-stochastic calculation. Furthermore, in order to be consistent with the analysis, we suppose that

[11] Hinich and Ordeshook, "Plurality Maximization Vs. Vote Maximization: A Spatial Analysis with Variable Participation," pp. 775-776.

(4.5) $g[x - \theta_1)'(x - \theta_1)] = g(||x - \theta_1||^2) = g(\xi^2)$

represents the <u>proportion</u> of citizens who are $\xi$ units from $\theta_1$, closer to $\theta_1$ than to $\theta_2$, and who abstain from voting.[12] The relevant probability measure, then, will simply be the relative frequency defined by (4.5). As in the previous discussion, we assume that if the $i^{th}$ citizen votes, it will be for the candidate at $\theta_1$ instead of the one at $\theta_2$ if $U_i(\theta_1) > U_i(\theta_2)$, and he will vote if and only if $U_i(\theta_1) > 0$. The relative frequency probability Pr* is a distribution on the space $\mathcal{U}$ of deterministic utility functions, and the probability that the $i^{th}$ citizen abstains if his favorite candidate is at $\theta_1$ is

(4.6) $Pr^*[U_i(\theta_1) \leq 0] = g(||x_i - \theta_1||^2)$.

Since $g$ is a function of the Euclidean distance from the voter's preferred point to his favorite candidate's platform vector (see equation (4.5)), it follows at once that if Pr* is defined in terms of relative frequencies, the proportion of citizens among the subset of the electorate a fixed distance $\xi$ from either $\theta_1$ or $\theta_2$ who abstain from voting is independent of the actual spatial location of $\theta_1$

[12] It is important to note that we are no longer utilizing the personal probability calculus of voting defined in (4.0). We are simply attempting to explore any conceivable rational interpretation of the unspecified portion of the H-O analysis. In particular, we are attempting to use (4.0) as a point of departure.

and $\theta_2$. This is easy to see by observing that

$$Pr^*[U_i(\theta_1) \le 0 \;\Big|\; ||x_i - \theta_1||^2 = \xi \le ||x_i - \theta_2||^2;\ \theta_1, \theta_2] = \gamma$$

where $\theta_1$ and $\theta_2$ are any two fixed candidate strategies, $x_i$ is on a "ring" a fixed radius $\xi$ from $V_i$'s favorite candidate, and $\gamma$, the proportion of citizens on the ring who abstain, is a constant (see Figure 4.1). But in order to be consistent with the Hinich-Ordeshook analysis, we must define the probability of abstaining as a function of distance to the closest candidate as in equation (4.6). Thus, the probability of abstaining, while a function of the distance to the closest candidate, is not a function of the <u>location</u> of the candidate.

Now consider the collection of citizens $x_i$ such that $||x_i - \theta_1||^2 = \xi \le ||x_i - \theta_2||^2$ for two fixed candidate strategies $\theta_1$ and $\theta_2$. Under this condition the probability that the $i^{th}$ citizen abstains is

$$Pr^*[U_i(\theta_1) \le 0] = Pr^*[f_i(||x_i - \theta_1||^2) \le 0]$$

$$= Pr^*[f_i^{-1}(0) \ge ||x_i - \theta_1||^2].$$

Hence,

$$(4.7) \qquad Pr^*[U_i(\theta_1) \le 0] = Pr^*[f_i^{-1}(0) \ge \xi].$$

Thus, given a particular pair of strategies, $\theta_1$ and $\theta_2$, the proportion of voters a fixed distance $\xi$ from $\theta_1$ and closer

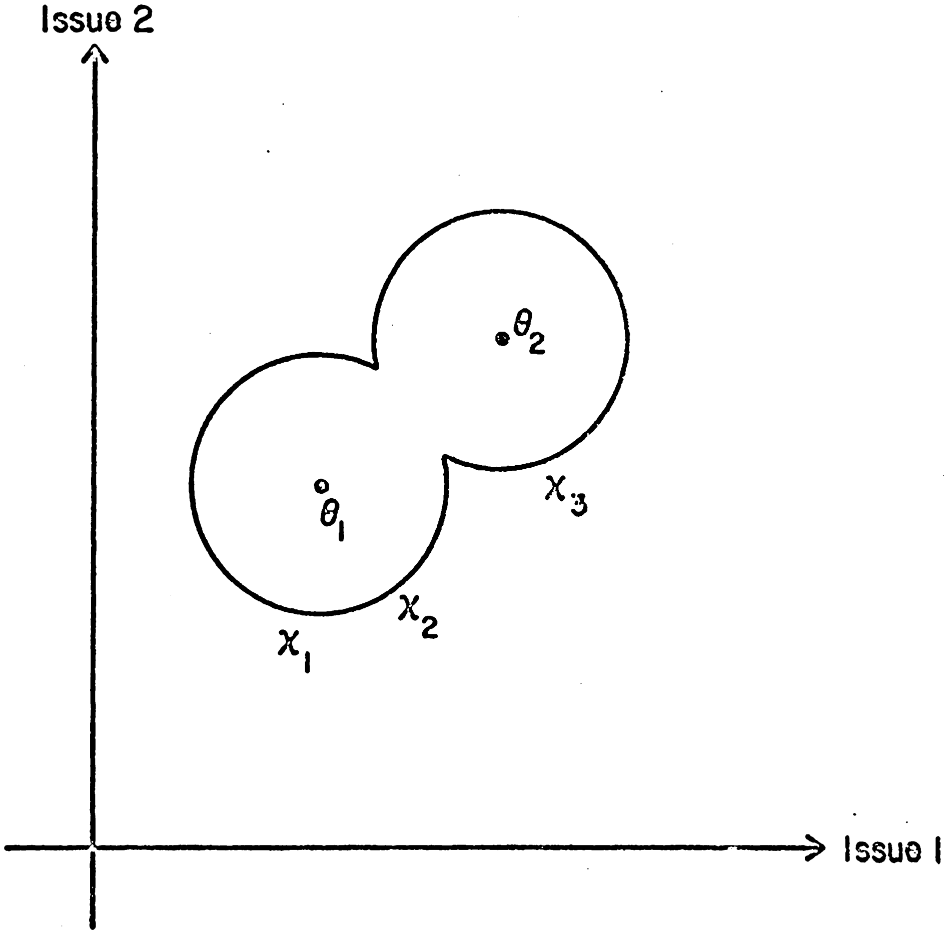



Figure 4.1

Three Voters in a Two-Dimensional Issue Space with Their Preferred Points Equidistant from Their Favorite Candidate's Strategy Vector to $\theta_1$ than $\theta_2$ who abstain is the proportion for which $f_i^{-1}(0) \geq \xi$. From (4.6) we see that $Pr^*[f_i^{-1}(0) \geq \xi] = g(\xi)$ does not depend on the specific locations of either $\theta_1$ or $\theta_2$; nor does it depend on the location of $x_i$ except insofar as $x_i$'s spatial position affects $||x_i - \theta_1||^2$ and $||x_i - \theta_2||^2$. From (4.7) we see that $Pr^*[f_i^{-1}(0) \geq \xi]$ is the distribution function of the random variable $f_i^{-1}(0)$.[13] Since $f_i^{-1}(0)$ is the precise distance which marks the citizen's threshold of abstaining (or voting), the fact that $f_i^{-1}(0)$ is a random variable stochastically independent of position means that the distribution of the distortion function $f_i$ must be independent of the $i^{th}$ citizen's preferred point $x_i$. Hence, we must assume that $f_i$ and, therefore, $U_i$, is independent of $x_i$.

The substantive interpretation of this independence is obvious. It indicates that the voters' spatial preferences and their utility functions must be uncorrelated. For example, citizens who are extremists with respect to spatial position cannot be extremists with respect to their actual utility functions. This means that the distribution of voters' utility functions at one social state must be, for

---

[13]The distribution function of the random variable $f_i^{-1}(0)$ is actually $Pr^*[f_i^{-1}(0) < \xi] = 1 - Pr^*[f_i^{-1}(0) \geq \xi]$. We could easily change the sense of our inequalities by discussing the probability of voting instead of the probability of abstaining.

the most part, similar to the distribution of the voters' utility functions at another social state--an obvious interpersonal comparison of utility.

The rather extensive amount of groundwork that we were required to lay prior to analyzing the probabilities discussed by Hinich and Ordeshook, as well as the detail of our dissertation, bespeak the subtle nature of their assumptive omission. Furthermore, we are not surprised that their formulation of abstentions due to alienation in a spatial model of voting under majority rule was not resolved as easily as appeared to be the case in their 1970 APSR paper. Our own experience with this problem is that each time we construct a mathematical model which incorporates abstentions due to alienation into the structure, we soon discover that we have also introduced interpersonal comparisons of utility into the model. We believe that it will ultimately be the case that a realistic model of political participation allowing abstentions but not admitting interpersonal comparisons will be formulated, but the mathematical structure necessary for such a formulation will be non-trivial. It is our prejudice that the problem of modeling optimal candidate strategies when citizens abstain is the outstanding one in spatial models of political participation today.

# CHAPTER V

## EQUIVALENCE OF CANDIDATES' OBJECTIVE FUNCTIONS

In a recent paper, Aranson, Hinich, and Ordeshook (henceforth denoted A-H-O) explored the nature of various election goals and strategies in a model of the electoral process.[1] In particular, they gave the usual definition of a spatial model, defined a spatial strategy, specified various candidate objective functions within the model, and formulated necessary and sufficient conditions for the objective functions to be equivalent. "Equivalence" is defined by characterizing an election as an n-person, zero-sum, non-cooperative, spatial game, and their theorems are constructed under certain assumptions about the expected proportions of the vote which the vying candidates receive under various combinations of spatial strategies and objective functions. No specific family of distributions on the outcomes of the election is assumed. While the A-H-O analysis is quite interesting, there is no doubt that the assumptions upon which their results are predicated reduce the attractiveness of their model, both as a representation of a "real-world" electoral process and as a model to be

[1] Peter Aranson, Melvin Hinich, and Peter Ordeshook, "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," American Political Science Review, Vol. 68 (March 1974), pp. 135-152.

taken seriously by candidates for office. To be more specific, we list the following set of assumptions, at least one of which undergirds each of the five theorems in the A-H-O analysis of equivalent candidate strategies:

i) if a candidate changes his strategy without affecting his expected proportion of the vote, then his opponents' expected proportions of the vote will remain unchanged.

ii) at a strategy equilibrium, every candidate's expected proportion of the vote is the same.

iii) "on the average," and prior to the election, each candidate is as likely to underestimate the magnitude of his vote as he is to overestimate it.

iv) the joint distribution of proportions of the vote for the candidates is multivariate normal.

While these assumptions do not hold in the "real world" and therefore seriously limit the usefulness of the A-H-O model, their paper is an important seminal work in the area of alternative candidate strategies.

In this chapter, we intend to address the question of equivalent strategies in a different manner. We construct a stochastic model which depends upon certain estimates made by the candidates and which specifies a joint distribution on the proportions of the vote for each of the candidates. More specifically, we assume that these proportions are Dirichlet random variables, and, thus, the stochastic element in a two-candidate election is completely characterized by three parameters.[2] We believe that the family of Dirichlet distributions is sufficiently versatile to model many natural phenomena; yet it demonstrates a degree of simplicity such that a candidate who is reasonably adept at estimating probabilities could easily use our model to make a fairly accurate estimate of the actual joint distribution of proportions of his and his opponents' vote for a fixed set of political strategies. Furthermore, our model is flexible in that it "allows" representation of a wide variety of electoral distributions without necessitating the restrictive assumptions of the A-H-O approach.

In the first section of this chapter, we concentrate on a two-candidate election and describe the relationship between "political platforms" and points in the parameter space of the Dirichlet distribution. Ultimately, within the context of this model we will prove the equivalence of the six candidate objective functions suggested by Aranson, et al., as well as a seventh which we append to their list.

---

[2] The family of Dirichlet distributions has been chosen for a variety of reasons to be presented in Section 5.3. An excellent discussion of numerous characteristics of these distributions may be found in Norman Johnson and Samuel Kotz, Distributions in Statistics: Continuous Multivariate Distributions, (New York: John Wiley & Sons, 1972), pp. 231-235.

## Section 5.1 Preliminaries

Although the model we develop is not a spatial model, there is a heuristic motivation for considering it from a spatial perspective. If the election entails a single race involving $p$ candidates, identified as $C_1, C_2, \ldots, C_p$, whose spatial strategies are $\theta_1, \theta_2, \ldots, \theta_p$ respectively, then we will call the vector $\underset{\sim}{\theta} = [\theta_1, \theta_2, \ldots, \theta_p]'$ a ballot, since it contains characterizations of each candidate's political platform and, thus, represents the voter's domain of choice for the race.

The citizen's input into the electoral calculus is his choice of one of the candidates. Formally, he selects a vector $c_i = [c_{i1}, c_{i2}, \ldots, c_{ip}]'$ such that $c_{ik}$ is either 0 or 1 and $\sum_{j=1}^{p} c_{ij} \leq 1$. The vector $c_i$, the $i^{th}$ citizen's vote, either contains all zeroes, indicating that he chooses to abstain from voting, or else it contains exactly one 1, thus specifying the candidate supported by the $i^{th}$ citizen. If there are $N$ citizens voting for $p$ candidates, then the totality of the vote in any election is an element in $\mathcal{W}$, the space of all $N \times p$ matrices with entries 0 or 1 and such that each row of the matrix sums to 0 or 1.[3]

---

[3]The $i^{th}$ row in $\mathcal{W}$ contains the vote of the $i^{th}$ citizen. If there are $k$ races in the election with $p_\beta$ candidates in the $\beta^{th}$ race, then the $i^{th}$ citizen's vote is a point in $\mathcal{W}'$, an appropriately specified $N \times (p_1 + p_2 + \ldots + p_k)$ matrix. Since there are $N$ voters, we need $N$ rows in the matrix to completely specify the vote.

We let $\phi_j$ denote the $j^{th}$ candidate's objective function, an algebraic expression whose substantive interpretation is the "payoff" received by the candidate if he adopts a certain strategy in the conduct of his political campaign. For example, if the $j^{th}$ candidate adopts the objective of maximizing the probability that his proportion of the vote is in excess of 60 percent of the total vote, and if the ballot is $\underset{\sim}{\theta}_0$, then the payoff is represented by $\phi_j(\underset{\sim}{\theta}_0)$. Once his objective function has been determined, the $j^{th}$ candidate may make comparative evaluations of accessible platforms simply by anticipating his opponents' strategies and contrasting the size of the payoff under various ballots. Thus, in this model, if the $j^{th}$ candidate adopts the objective function $\phi_j$, he prefers $\underset{\sim}{\theta}$ over $\underset{\sim}{\theta}'$ if and only if $\phi_j(\underset{\sim}{\theta}) > \phi_j(\underset{\sim}{\theta}')$, and he is indifferent between these ballots if and only if $\phi_j(\underset{\sim}{\theta}) = \phi_j(\underset{\sim}{\theta}')$.

The functions $\phi_j$ must be "interpreted" rather carefully. In particular, the reader should go to some lengths to avoid associating the objective functions with utility functions, since, in this comprehensive model, any reasonable utility function would take into account <u>both</u> the ballot of strategies and the vector of objective functions which the candidates adopt. In other words, the $j^{th}$ candidate's objective function is a mapping $\phi_j : S \to E$, where $S = \{(\theta_1, \theta_2, \ldots, \theta_p) | \theta_j$ is the $j^{th}$ candidate's spatial strategy$\}$ and $E$ is the set of real numbers; whereas an appropriate utility function for the $j^{th}$ candidate would be a mapping $U_j : \mathcal{O}\times\mathcal{S} \to \mathcal{E}$, where $\mathcal{O} = \{(\phi_1,\phi_2,\ldots,\phi_p) | \phi_j$ is the $j^{th}$ candidate's objective function$\}$. Values of the objective functions should be interpreted to mean something like "voter support," e.g., the number of votes for a candidate or the candidate's proportion of the total vote. To illustrate further, it could easily be the case that if the $j^{th}$ candidate's utility were ascertainable at all, $\phi_j(\underset{\sim}{\theta}) > \phi_j'(\underset{\sim}{\theta})$ even though $U_j(\phi_j,\underset{\sim}{\theta}) < U_j(\phi_j',\underset{\sim}{\theta})$. If $\phi_j$ represents the candidate's objective of "maximizing the probability that his vote exceeds a particular level" and $\phi_j'$ represents his objective of "maximizing the probability that his plurality exceeds some level," then it may be that his payoff (the actual probability) is greater for the former while his utility is greater for the latter.

To summarize and consolidate these ideas, we suggest a rather rudimentary, formal definition of an election. An election involving a single race in which $p$ candidates compete with each other for the votes of an electorate consisting of $N$ citizens and in which $n$ "relevant" issues can be identified can be thought of as a point in the space $\mathcal{O}\times\mathcal{S}\times\mathcal{W}$.[4] In other words, the essential ingredients of any

---

[4]Since many elections entail more than one race, an election is actually a point in the space $\mathcal{O}'\times\mathcal{S}'\times\mathcal{W}'$, where $\mathcal{O}'$ and $\mathcal{S}'$ are simply $\mathcal{O}$ and $\mathcal{S}$ expanded to include the objective functions and spatial strategies respectively of all of the $\sum_{\beta=1}^{k} p_\beta$ candidates and $\mathcal{W}'$ is defined in footnote 3.

election consist of the general goals and orientations of the candidates (formalized by their objective functions); their announced, or otherwise publicly known, positions on the issues of the campaign; and, ultimately, the response of the electorate to the candidates for office (as manifested in their votes).

Now, given a particular ballot $\underset{\sim}{\theta}_0$, let $V_j(\underset{\sim}{\theta}_0)$ for $j=1,2,\ldots,p$ be the proportion of the eligible voters who vote for the $j^{th}$ candidate. Since an eligible voter may choose to write in the name of an individual who is not on the ballot or, perhaps, abstain from voting altogether, it follows that $\sum_{j=1}^{p} V_j(\underset{\sim}{\theta}_0) \leq 1$. Notice that each $V_j$ is a function of the ballot $\underset{\sim}{\theta}_0 = [\theta_1,\theta_2,\ldots,\theta_p]'$, that is, a function of <u>all</u> of the candidates' strategies. Furthermore, we assume that each $V_j$ is a random variable, the "randomness" being a function of the candidate's uncertainty about the outcome of the election. If each $V_j$ were known, and therefore not random, then the election would be completely determined and the candidate's choice of an objective function would be purely academic.

Although the model could be formulated in a manner that requires each candidate to make direct (point) estimates of $V_1(\underset{\sim}{\theta}),V_2(\underset{\sim}{\theta}),\ldots,V_p(\underset{\sim}{\theta})$, it is improbable that he could do so with any reasonable degree of accuracy, and it follows that a model based upon point estimates would probably be misleading to a candidate attempting to utilize it to guide his choice of a strategy under a particular objective function. Consequently, we favor a model which requires the candidate to estimate the distributions of the random variables $V_1(\underset{\sim}{\theta}), V_2(\underset{\sim}{\theta}), \ldots, V_p(\underset{\sim}{\theta})$ and then treat the payoff under the objective function $\phi_j$ as a function of the joint distribution of the proportions of the vote each candidate will receive. Thus, we focus attention on the joint distribution of $V_1(\underset{\sim}{\theta}), V_2(\underset{\sim}{\theta}), \ldots, V_p(\underset{\sim}{\theta})$.

The relevant objective functions are displayed in Table 5.1, where we use $\nu_j(\underset{\sim}{\theta})$ to represent the expected value $E[V_j(\underset{\sim}{\theta})]$ of $V_j(\underset{\sim}{\theta})$. Our model assumes that the $j^{th}$ candidate, after adopting one of the seven objective functions, say $\phi_j$, as his modus operandi, attempts to locate a spatial strategy $\theta_j$ for himself and otherwise affect the remaining p-1 candidates' choices of platforms $\theta_k$, $k \neq j$, so that $\phi_j(\underset{\sim}{\theta})$ is a maximum, where $\underset{\sim}{\theta} = [\theta_1, \theta_2, \ldots, \theta_p]'$. In order to facilitate reference, we have listed (and numbered) the first six objective functions in a manner consistent with the ordering in the A-H-O paper.[5] A rather concise description of each of these functions is included in their paper, and they make an effort to contrast the relative merits of each.

[5] Aranson, Hinich, and Ordeshook, "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," p. 139.

Table 5.1

The j[th] Candidate's Objective Functions[6]

| Function | Substantive Interpretation |
|---|---|
| $O_1$: $\phi_j = E[V_j(\underset{\sim}{\theta}) - \max_{k \neq j}\{V_k(\underset{\sim}{\theta})\}]$ | Expected Plurality |
| $O_2$: $\phi_j = \nu_j(\underset{\sim}{\theta}) / \sum_{k=1}^{p} \nu_k(\underset{\sim}{\theta})$ | Proportion of the Expected Votes |
| $O_3$: $\phi_j = \nu_j(\underset{\sim}{\theta})$ | Expected Vote |
| $O_4$: $\phi_j = Pr[V_j(\underset{\sim}{\theta}) - \max\{V_k(\underset{\sim}{\theta})\} \geq \lambda_j]$ | Probability that Plurality Exceeds Some Level |
| $O_5$: $\phi_j = Pr[V_j(\underset{\sim}{\theta}) / \sum_{k=1}^{p} V_k(\underset{\sim}{\theta}) \geq \lambda_j]$ | Probability that Proportion Exceeds Some Level |
| $O_6$: $\phi_j = Pr[V_j(\underset{\sim}{\theta}) \geq \lambda_j]$ | Probability that Vote Exceeds Some Level |
| $O_7$: $\phi_j = E[V_j(\underset{\sim}{\theta}) / \sum_{k=1}^{p} V_j(\underset{\sim}{\theta})]$ | Expected Proportion of the Vote |

[6] In "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," Aranson, Hinich, and Ordeshook misspecify the first objective function, expected plurality, as

$$\phi_j = \nu_j(\underset{\sim}{\theta}) - \max_{k \neq j}\{\nu_k(\underset{\sim}{\theta})\}.$$

The correct formulation for $O_1$ is

$$E[V_j(\underset{\sim}{\theta}) - \max_{k \neq j}\{V_k(\underset{\sim}{\theta})\}],$$

and the two are equal only when

$$E[\max_{k \neq j}\{V_k(\underset{\sim}{\theta})\}] = \max_{k \neq j}[E(V_k(\underset{\sim}{\theta}))].$$

In addition, several functions must be multiplied by N in order to make the substantive interpretation consistent with the functional form. For example, the j[th] candidate's <u>expected vote</u> is $\phi_j = N\nu_j(\underset{\sim}{\theta})$.

We tend to discount the importance of $O_2$, since in our model it is more difficult for the candidate to estimate it than $O_7$, the expected proportion of the vote. Notice that in order for a candidate to estimate $O_2$, he must estimate the expected values of the distributions of every candidate's proportion of the vote.[7] On the other hand, in order to estimate the payoff under $O_7$, the candidate is only required to estimate the expected value of his proportion of the vote. It is interesting that the function $O_7 \left(\phi_j = E[V_j(\underset{\sim}{\theta})/\sum_{k=1}^{p} V_k(\underset{\sim}{\theta})]\right)$ is mathematically less tractable than $O_2 \left(\phi_j = E[V_j(\underset{\sim}{\theta})/\sum_{k=1}^{p} E(V_k(\underset{\sim}{\theta}))]\right)$ even though it is probably easier for the candidate to estimate. Furthermore, we claim that the distribution of a candidate's proportion of the vote is a rather natural distribution for the candidate to want to evaluate, and the mean of this distribution is probably the simplest parameter for him to <u>accurately</u> estimate.

---

[7]Since $O_2$ is essentially defined by $\phi_j \underset{\sim}{\theta} = E[V_j(\underset{\sim}{\theta})/\sum_{k=1}^{p} E\left(V_k(\underset{\sim}{\theta})\right)]$, any candidate attempting to discern between attainable strategies by estimating the ultimate payoff with respect to each social state would be forced to estimate $E[V_k(\underset{\sim}{\theta})]$ for $k = 1,2,\ldots,p$. It could be argued that if the $j^{th}$ candidate could predict the expected proportion of the electorate which will abstain, then the denominator of $O_2$ is 1 minus that number. Unfortunately, it is doubtful that this estimate could be made without first estimating the expected proportions of the citizenry who will vote for each of the p candidates.

We will use the usual definition of an equilibrium ballot, taking into account the fact that candidates may differ in terms of both objective and strategy.

<u>Definition 5.1</u>: The vector $\underset{\sim}{\theta}^* = [\theta_1^*, \theta_2^*, \ldots, \theta_p^*]'$ is an <u>equilibrium ballot</u> with respect to $[\phi_1, \phi_2, \ldots, \phi_p]'$ if and only if $\phi_j(\theta_1^*, \theta_2^*, \ldots, \theta_j^*, \ldots, \theta_p^*) \geq \phi_j(\theta_1^*, \theta_2^*, \ldots, \theta_j, \ldots, \theta_p^*)$ for all $\theta_j$ and for $j = 1, 2, \ldots, p$.

Less formally, $\underset{\sim}{\theta}^*$ is an equilibrium point if no candidate will find it to his advantage to relocate his strategy provided none of his opponents simultaneously "move away" from their positions. Therefore, if we assume that collusion among candidates is not allowed (or is otherwise disadvantageous), equilibrium strategies become very attractive ones for the candidates.[8] Another concept which is quite useful is that of a weak equilibrium.

<u>Definition 5.2</u>: A ballot $\underset{\sim}{\hat{\theta}} = [\theta_1, \theta_2, \ldots, \hat{\theta}_j, \ldots, \theta_p]'$ is a <u>weak equilibrium point</u> with respect to $\phi_j$ if and only if $\phi_j(\theta_1, \theta_2, \ldots, \hat{\theta}_j, \ldots, \theta_p) \geq \phi_j(\theta_1, \theta_2, \ldots, \theta_j, \ldots, \theta_p)$ for all $\theta_j$ and for a fixed choice of $\theta_1, \theta_2, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_p$.

In other words, $\underset{\sim}{\hat{\theta}}$ is a weak equilibrium with respect to $\phi_j$ if, given the strategy choices of all other candidates, $C_j$

---

[8] R. Duncan Luce and Howard Raiffa, <u>Games and Decisions</u>, (New York: John Wiley and Sons, 1957), pp. 106-107, 170-171.

cannot improve his position by unilaterally altering his strategy. A weak equilibrium for $C_j$ is conditioned on the political platforms of $C_j$'s opponents, while a strong equilibrium is, in that sense, unconditional. If a ballot is simultaneously a weak equilibrium for every candidate for every possible choice of strategies of the remaining candidates, with $[\phi_1,\phi_2,\ldots,\phi_p]'$ fixed, then that ballot is a strong equilibrium with respect to $[\phi_1,\phi_2,\ldots,\phi_p]'$.

We adopt the A-H-O definition of equivalent objective functions.

<u>Definition 5.3</u>: Two objective functions $\phi_j$ and $\phi_j'$ are <u>equivalent</u> for $V_j$ if whenever $\underset{\sim}{\theta}*$ is an equilibrium with respect to $[\phi_1,\phi_2,\ldots,\phi_j\ldots,\phi_p]'$, it follows that $\underset{\sim}{\theta}*$ is an equilibrium with respect to $[\phi_1,\phi_2,\ldots,\phi_j',\ldots,\phi_p]'$.[9]

Assume for the moment that all of the p candidates in a particular race have chosen their objective functions and that C is a non-empty set of all spatial equilibria with respect to $[\phi_1,\phi_2,\ldots,\phi_j,\ldots,\phi_p]'$.[10] Then equivalent objec-

[9] It is also possible to define <u>weak-equivalent objective functions</u> by specifying that $\phi_j$ and $\phi_j'$ are weak-equivalent for the $j^{th}$ candidate if whenever $\hat{\underset{\sim}{\theta}}$ is a weak equilibrium point with respect to $[\phi_1,\phi_2,\ldots,\phi_j,\ldots,\phi_p]'$, it follows that $\hat{\underset{\sim}{\theta}}$ is a weak equilibrium with respect to $[\phi_1,\phi_2,\ldots,\phi_j',\ldots,\phi_p]'$.

[10] It may be that an election has no equilibrium strategies, i.e., that C is the empty set. It is vacuously true that in such a situation all objective functions are equivalent.

tive functions may be understood by imagining an election which is identical in all respects to the initial race except that $C_j$ now has $\phi_j'$ as his objective function. "Identical in all respects" means that, in addition to all other candidates adopting the same objective functions that they employed in the first case and having the same issue space from which to choose their strategies, precisely the same set $C$ of equilibrium points eventuates.

It is clear that it would be useful to know whether equilibrium strategies are invariant with respect to different choices of objective functions in the sense that a candidate who is attempting to maximize the objective function $\phi_j$ will adopt the same strategy that he would if he were attempting to maximize $\phi_j'$, all other things being equal. For example, should a candidate whose goal is to maximize his vote express commitment to the same political platform that he would embrace if he were attempting to maximize his plurality? In the next section we present the details of a model which we believe allows a reasonably faithful representation of the candidate's decision process, and then in Section 5.3 we show that in our model the seven objective functions listed in Table 5.1 are equivalent.

## Section 5.2 The Model

In their paper, Aranson, et al. define the concepts of symmetric and strongly symmetric elections, and two of their major results (Theorems 2 and 3) are based upon the assumption that the election in question is one of these two types.[11] As we have previously indicated, it is not likely that models which posit symmetric or strongly symmetric elections can represent "real-world" phenomena. Furthermore, one of the A-H-O results (Theorem 4) is based on the assumption that the $j^{th}$ candidate estimates the true proportion of his vote $V_j(\underset{\sim}{\theta})$ to be $\nu_j(\underset{\sim}{\theta})$, with the true and estimated proportions related by the additive model $V_j(\underset{\sim}{\theta}) = \nu_j(\underset{\sim}{\theta}) + \varepsilon_j$, where, for $j = 1,2,\ldots,p$, $\varepsilon_j$ is a randomly distributed prediction error whose distribution is independent of $\underset{\sim}{\theta}$ and such that $E[\varepsilon_j] = 0$.[12] This model is not inconsistent with our perception of what a candidate may actually do, i.e., he may evaluate the desirability of various strategies by estimating and contrasting his own proportion of the vote at each of several platforms, or he may estimate both his and his opponents' proportions of the vote and compute some function of those estimates. On the other hand, we have previously noted that it is probably

[11]Aranson, Hinich, and Ordeshook, "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," pp. 137-138.

[12]It is noteworthy that with this fairly weak assumption, $O_1$ is equivalent to $O_4$, $O_3$ is equivalent to $O_6$, and, hence, maximizing expected plurality and maximizing the probability of winning in a plurality structured election are equivalent.

unrealistic to believe that candidates do or could make accurate point estimates of these proportions, and it is most unlikely that the prediction error is random and independent of the platforms.[13] Furthermore, the A-H-O model does not provide the candidate with a method for specifying the uncertainty associated with his estimate, i.e., he has no formal mechanism which enables him to estimate the variance of the error. A more realistic model, and one which is widely used in the area of probability estimation, assumes that a candidate is capable of estimating the probability that his proportion of the vote will fall within certain limits, e.g., he may believe that the probability is 0.8 that his proportion of the vote will be between 0.4 and 0.6. Of course, the probability specification in such a statement is subjective; nevertheless, with several subjective estimates of this sort, a model based upon the family of beta distributions may provide a good approximation to the distributions of the candidates' proportions of the vote, with one univariate distribution for each of the $p$ candidates. Furthermore, if these distributions are indexed by (or functions of) ballots of candidates' platforms, then each candidate could make a rational choice among strategies as

[13]Theoretically, it is absurd to assume that the error has mean 0 and is independent of the proportion of the vote, since this would allow the possibility of negative estimates.

a function of how such a choice would ultimately affect his estimated distribution of the proportion of the vote. To be more precise, since it is the entire ballot, and not just the $j^{th}$ candidate's platform, which determines the proportion of the vote which goes to each candidate, a realistic model should include a p-dimensional joint distribution of the candidates' proportions of the vote. It would be convenient if the marginal distributions of this joint distribution were members of the previously mentioned beta family of univariate distributions. In addition, we are concerned with more than just the proportions of the vote which go to the various candidates; we are actually interested in certain functions of those proportions.[14] Finally, the model should be formulated from the candidate's perspective, since it is the candidate who must estimate the parameters of the election in order to choose a strategy.

In reality, we are attempting to model a first-order approximation to a candidate's thought process when he evaluates accessible social states for the purpose of choosing one which is likely to maximize a given objective

---

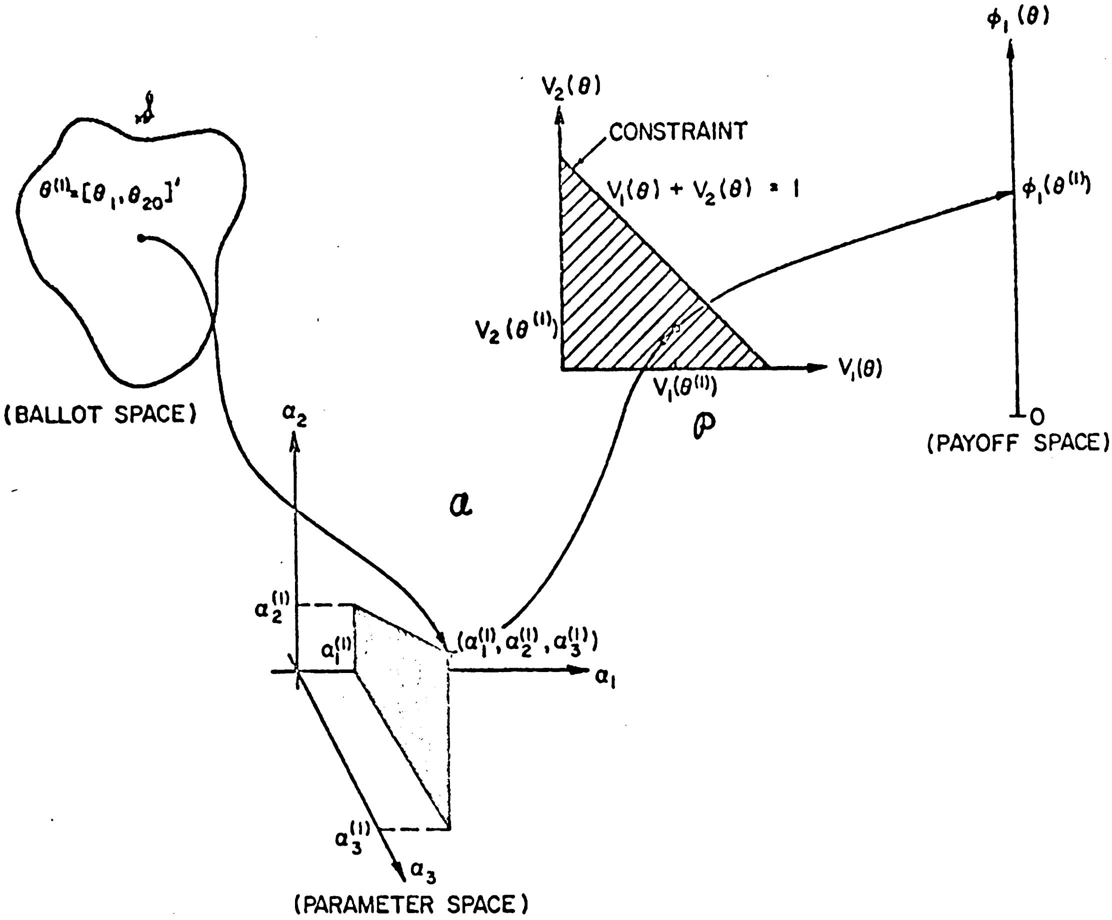
[14]By way of example, if the proportions of the vote going to the $p$ candidates are $V_1(\underset{\sim}{\theta}), V_2(\underset{\sim}{\theta}), \ldots, V_p(\underset{\sim}{\theta})$ respectively for a given ballot $\underset{\sim}{\theta}$, then the $j^{th}$ candidate $C_j$ may be interested in

$$\phi_j\left(V_1(\underset{\sim}{\theta}), V_2(\underset{\sim}{\theta}), \ldots, V_p(\underset{\sim}{\theta})\right) = E[V_j(\underset{\sim}{\theta}) / \sum_{k=1}^{p} V_k(\underset{\sim}{\theta})],$$

his expected proportion of the vote.

function. The model which we introduce is fairly sophisticated, yet it leads us to the conclusion that all of the objective functions in Table 5.1 are equivalent. This result is negative in the sense that we do not believe that candidates perceive all optimal strategies under one objective function as necessarily optimal under another. Consequently, we conclude that the subtleties of different elections, insofar as determining dominant strategies is concerned, may be beyond the scope of any reasonably accurate and tractable formal model.

The ideas presented above are summarized in Figure 5.1, where, for the purpose of constructing a display, we confine our analysis to a two-candidate election, and, without loss of generality, evaluate the situation from the point of reference of a first candidate, $C_1$. We will assume that $C_1$ is cognizant of the fact that there is a variety of election objectives from which he may choose, and that he has decided to attempt to maximize the objective function $\phi_1$. Having made that decision, he must now select a strategy (political platform) from a subspace of the issue space which is at least implicitly constrained in two ways. First, it is probably the case that his platform must be within fairly close proximity to his actual preference on the




<u>Figure 5.1</u>

Structural Components of

the Candidate's Decision Mechanism

issues.[15] Specification of this subspace is, for the most part, a decision which reflects the nature of the candidate's political ideology; consequently, the second candidate's announced (or otherwise known) position on the issues has very little affect on this phase of $C_1$'s choice. Second, among the platforms in the previously identified subspace, $C_1$ should (if he is rational in a mathematical sense) select a position which maximizes his objective function. Since $C_1$'s payoff with respect to his objective function is a function not only of his own platform, but also of $C_2$'s spatial strategy, it follows that his choice should actually be made from a subset $S_1$ of the space $S = \{\underset{\sim}{\theta}=[\theta_1, \theta_2]' \mid \theta_j \varepsilon E^n\}$ of ballots. In other words, $C_1$ must have as input for his decision an anticipation of $C_2$'s behavior. We are not so pessimistic that we believe that a candidate would endorse _any_ position if he thought that it would maximize his objective function. On the other hand, we are not so unrealistic that we fail to recognize that a candidate's announced position on the issues may deviate from that point which actually represents his political ideology, that deviation being a function of both his estimates of his opponents'

---

[15] This is true for several reasons. First, no one can completely ignore his preferences. Second, candidates cannot suddenly alter their platforms without casting doubt on their ideologic stability and leadership ability. Third, an office won by effecting political trade-offs inconsistent with one's preferences may not be worth the price of victory.

platforms and his estimate of the distribution of preferences and loss functions of the electorate.

In our model we assume that there is a family of bivariate density functions $f[\cdot,\cdot|\underset{\sim}{\theta}]$ defined on the space $P=\{(V_1(\underset{\sim}{\theta}),V_2(\underset{\sim}{\theta}))|V_1(\underset{\sim}{\theta}) \geq 0,\ V_2(\underset{\sim}{\theta}) \geq 0, \text{ and } V_1(\underset{\sim}{\theta})+V_2(\underset{\sim}{\theta}) \leq 1\}$ of proportions of the vote going to each of the two candidates. In the next section we will specify a family of distributions which can be realistically utilized in this analysis, but for the moment we shall ignore such considerations. Notice, however, that each member $f[\cdot,\cdot|\underset{\sim}{\theta}]$ of the family of distributions is dependent on the ballot $\underset{\sim}{\theta}$.

We imagine that the mechanism with which the first candidate selects his platform is the following:

(1) First, he anticipates $C_2$'s political platform with as much accuracy as possible.

(2) Then with his estimate of $C_2$'s strategy $\theta_{20}$ in mind, he searches among those ballots $\underset{\sim}{\theta}^{(1)} = [\theta_1,\theta_{20}]'$ in $S_1$ for the one which produces a joint distribution $f[\cdot,\cdot|\underset{\sim}{\theta}^{(1)}]$ on $P$ which ultimately has the effect of maximizing his objective function $\phi_1$ if, indeed, $C_2$ actually does choose $\theta_{20}$.

(3) With each consideration of $\underset{\sim}{\theta}$ in step (2), $C_1$ must estimate the proportion of the electorate which will abstain from voting if, based upon his and $C_2$'s choices of strategies, $\underset{\sim}{\theta}$ is the ballot which ultimately eventuates.

Notice that $C_1$ does not necessarily want that ballot $\underset{\sim}{\theta}^{(1)}$ which "produces" the distribution which maximizes his proportion of the vote $V_1(\underset{\sim}{\theta}^{(1)})$ given that $C_2$ chooses $\theta_{20}$; for such a choice may give $C_2$ a large strategic advantage with respect to the objective function which $C_1$ has adopted. In short, $C_1$ may not even be interested in his proportion of the vote; for example, he may be much more concerned with maximizing his expected plurality, $E[V_1(\underset{\sim}{\theta}) - V_2(\underset{\sim}{\theta})]$. That is why the mapping from $\mathcal{P}$ to the payoff space is an important characteristic of this model (see Figure 5.1). It is noteworthy that for a given family of distributions on $\mathcal{P}$ the first candidate is actually implicitly specifying parameters (and, thus, identifying one member of the family of distributions) when he estimates his opponent's platform, guesses what proportion of the electorate will abstain from voting, and then chooses his own strategy. In other words, the first candidate defines--at least implicitly--a function $\alpha_1$ from a subset $S_1$ of the ballot space $\mathcal{S}$ <u>onto</u> a subset $A_1$ of the parameter space $\mathcal{A}$ such that

$\alpha_1(\theta_1,\theta_{20}) = (\alpha_1^{(1)},\alpha_2^{(1)},\alpha_3^{(1)})$.[16] In essence, in our model $C_1$ is required to estimate various parameters of the election, and $\alpha_1$ is the function which he uses to do so. Furthermore, even in a two-candidate election, $C_1$ may be forced to estimate more than two parameters, so $\alpha_1$ may actually be a function onto a range of triples (or, more generally, k-tuples) of parameters. The parameter space itself is the first (positive) octant of $E^3$, so the range of $A_1$ of $\alpha_1$, whether it is a bounded set or not, is a set of triples of positive numbers. If $A_1$ (or $A_2$ for that matter) is unbounded, we will assume that it is a Cartesian product of intervals or rays, for example,

$A_1 = \{(\alpha_1^{(1)},\alpha_2^{(1)},\alpha_3^{(1)}) \mid \alpha_1^{(1)} > \alpha_{10},\alpha_2^{(1)} > \alpha_{20},\alpha_{30} < \alpha_3^{(1)} < \alpha'_{30}\}$.

While the theorems in Section 5.3 are true in a more general setting, we see no practical justification for considering arbitrary unbounded subsets of the parameter space.

Since we are basically concerned with the equivalence (or nonequivalence) of a candidate's possible objective functions, it is important to understand the nature of an

---

[16] We have taken the liberty of using the same symbol $\alpha_1$ for $C_1$'s decision function and for the first component of the image of a ballot under that function.

Estimating the parameters $(\alpha_1^{(1)},\alpha_2^{(1)},\alpha_3^{(1)})$ as a function of the ballot relates our model to the spatial model. In general, however, our model, while it requires the candidate to estimate parameters, does not require him to utilize a spatial formulation.

equilibrium ballot in this model. The decision mechanism has thus far been described from the perspective of the first candidate, $C_1$. Naturally, $C_2$ is making similar calculations; he, too, has a function $\alpha_2$ from $S_2 \subset S$ onto $A_2 \subset A$, and his selection of a ballot $\underset{\sim}{\theta}^{(2)} = [\theta_{10},\theta_2]'$, where $\theta_{10}$, $C_2$'s estimate of $C_1$'s platform, is for the purpose of effecting a distribution on $P$ which will maximize his objective function $\phi_2$, whatever that function may be. If, in making these decisions, $C_1$'s choice of a ballot is $\underset{\sim}{\theta}^{(1)} = [\theta_1{}^*,\theta_{20}]'$ and $C_2$'s choice of a ballot is $\underset{\sim}{\theta}^{(2)} = [\theta_{10},\theta_2{}^*]'$, and if

i) $\theta_{10} = \theta_1{}^*$ and $\theta_{20} = \theta_2{}^*$

ii) $\phi_1([\theta_1{}^*,\theta_{20}]') \geq \phi_1([\theta_1,\theta_{20}]')$ for all $\theta_1$ and
$\phi_2([\theta_{10},\theta_2{}^*]') \geq \phi_2([\theta_{10},\theta_2]')$ for all $\theta_2$,

then the ballot $\underset{\sim}{\theta}^* = [\theta_1{}^*,\theta_2{}^*]'$ is an <u>equilibrium</u>. In other words, if the two candidates choose ballots in such a manner that their estimates of their opponents' platforms actually coincide with those platforms and such that, once that choice is made, neither candidate can increase his payoff by unilaterally altering his spatial strategy, then the common ballot $\underset{\sim}{\theta}^*$ adopted by the two candidates is an equilibiium point.[17]

[17] It is a trivial matter to extend the definition of an equilibrium strategy to a p-candidate election.

We will summarize the structure of our model by contrasting it with the A-H-O model. In "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives," Aranson, et al. (implicitly) utilized a function which maps ballots $\underset{\sim}{\theta}$ into payoffs $\phi_j(\underset{\sim}{\theta})$ without specifying either the parameter space $A$ or the space $P$ on which the family of distributions is defined. Such an analysis ignores several important characteristics of the decision-making process, and, in particular, it sheds little light on the selection mechanism which links a candidate's choice of a ballot to his payoff with respect to a given objective function. Furthermore, since essential features of the decision process have been omitted, the assumptions which underlie the theorems in the A-H-O analysis are, for the most part, unrealistic. In the next section, we will attempt to support the formulation of our model by filling in some details of the structure outlined in this section. In addition, we will show that in our model, no matter which of the seven objective functions listed in Table 5.1 is adopted by a candidate, the set $C$ of equilibria is invariant over elections, i.e., <u>all</u> of the objective functions are equivalent. Again, we feel that this result does not imply that these strategies are equivalent in the "real world"; it may indicate that assessing differences between strategies for different objective functions is beyond the scope of formal analysis by an actual candidate. In fact, we doubt that a realistic formal model can be developed which will guide a candidate making a choice between attainable platforms under such conditions, and our analysis in the next section provides a basis for our prejudice.

## Section 5.3 Equivalence of Objective Functions in a Two-Candidate Election

Ideally, one would like to complete the model poposed in the last section by assuming that any continuous bivariate distribution defined on $\mathcal{P}$ may be used to represent the joint distribution of the proportions of the vote as a function of the candidates' determination of a ballot. Unfortunately, such a formulation would not provide us with enough structure to investigate the relevant questions pertaining to equivalence of objectives, or else it would be too poorly defined for an actual candidate for office to use in his decision process. The purpose of the distributions on $\mathcal{P}$ is to answer such questions as, "If $C_1$ and $C_2$ endorse platforms $\theta_1$ and $\theta_2$ respectively, what is the probability that the first candidate's proportion of the vote is between forty and sixty percent?" In order to formulate answers for questions such as this and have a reasonable degree of confidence in the accuracy of the answer, it is imperative that we choose a versatile, yet simple, <u>family</u> of distributions defined on the space $\mathcal{P}$ such that a specific family member is a function of $\underset{\sim}{\theta} = [\theta_1, \theta_2]'$. One family of distributions with many characteristics which are desirable for this analysis, a family which is widely used in the analysis of probability estimation, is the class of Dirichlet distributions defined on a two-dimensional space.

If we let $V_1$ and $V_2$ be abbreviations for the random variables $V_1(\underset{\sim}{\theta})$ and $V_2(\underset{\sim}{\theta})$, then the density function for a two-dimensional Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \alpha_3)$ is

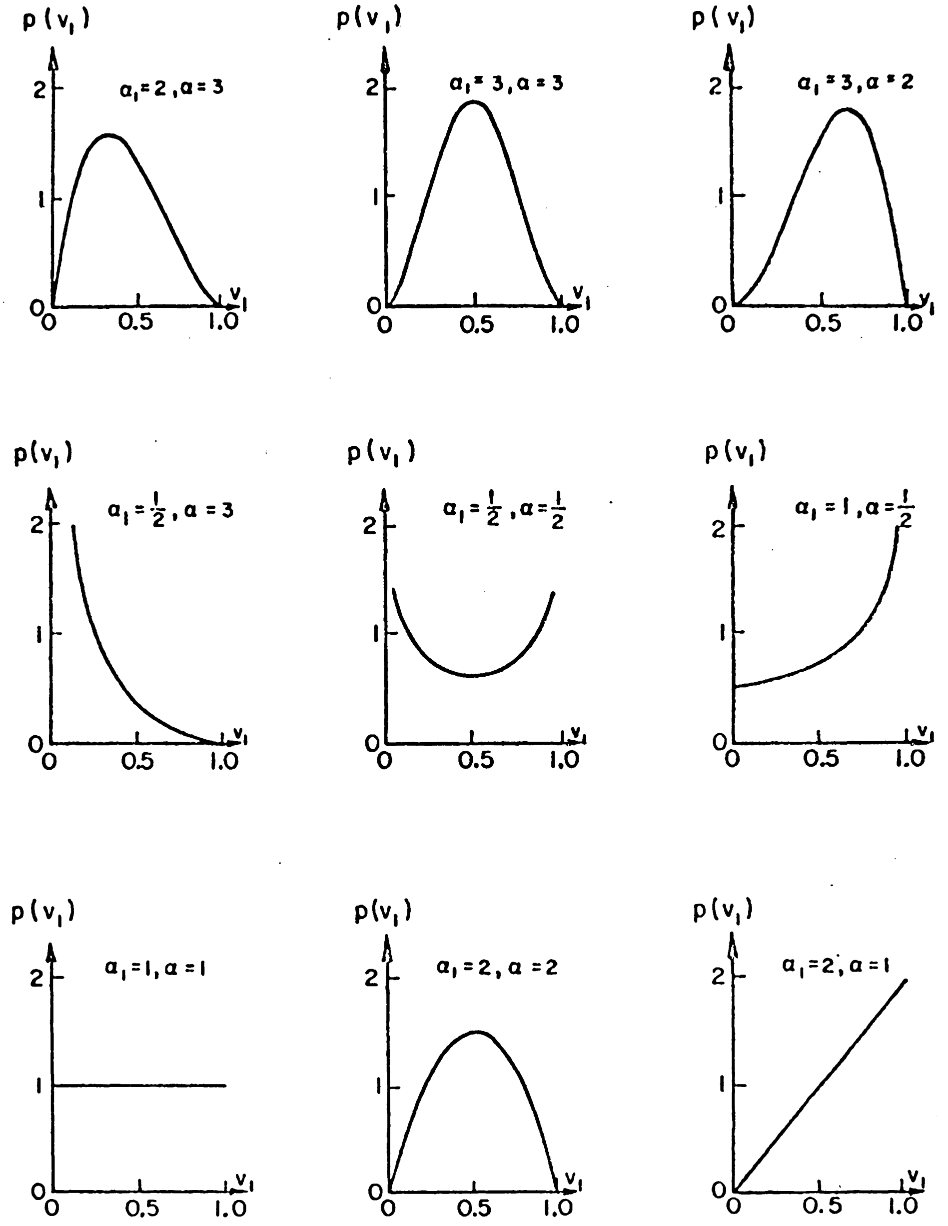
$$p_{V_1 V_2}(v_1, v_2) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} v_1^{\alpha_1 - 1} v_2^{\alpha_2 - 1} (1 - v_1 - v_2)^{\alpha_3 - 1},$$

where $v_j \geq 0$ for $j = 1,2$ and $v_1 + v_2 \leq 1$. Among the desirable characteristics of a Dirichlet distribution is the fact that its marginal distributions belong to the class of beta distributions, a family which has considerable versatility for fitting univariate data (see Figure 5.2).[18] In particular, we have:

$$V_1 \sim \beta(\alpha_1, \alpha), \text{ where } \alpha = \alpha_2 + \alpha_3.$$

$$V_2 \sim \beta(\alpha_2, \alpha'), \text{ where } \alpha' = \alpha_1 + \alpha_3.$$

[18] Norman L. Johnson and Samuel Kotz, Distributions in Statistics: Continuous Univariate Distributions - 2, (New York: John Wiley and Sons, 1970), pp. 37-56.




<u>Figure 5.2</u>

Examples of Beta Distributions

It can be shown that the mean and variance of $V_1$ are

$$\nu_1(\underset{\sim}{\theta}) = E(V_1) = \alpha_1/(\alpha_1 + \alpha_2 + \alpha_3) \quad \text{and}$$

$$\sigma_{V_1}^2 = \mathrm{Var}(V_1) = \frac{\alpha_1(\alpha_2 + \alpha_3)}{(\alpha_1 + \alpha_2 + \alpha_3)^2 + (\alpha_1 + \alpha_2 + \alpha_3 + 1)}.$$

Of course, the mean and variance of $V_2$ have similar functional forms.

The Dirichlet distribution is ideally suited for this analysis since the relevant random variables in the voting model are <u>proportions</u> of the electorate which support the various candidates or else abstain from voting, and the Dirichlet distribution is a joint distribution of a random pair $(V_1,V_2)$ such that $v_j \geq 0$ for $j = 1,2$ and $v_1 + v_2 \leq 1$. For any $V_1$ and $V_2$, $1 - V_1 - V_2$ may be interpreted to be the proportion of the citizenry which abstains from voting. If we define a random variable $V_3 = 1 - V_1 - V_2$, then

$$V_3 \sim \beta(\alpha_3,\alpha''), \text{ where } \alpha'' = \alpha_1 + \alpha_2.$$

Of course, it follows that

$$\nu_3(\underset{\sim}{\theta}) = E(V_3) = \alpha_3/(\alpha_1 + \alpha_2 + \alpha_3) \quad \text{and}$$

$$\sigma_{V_3}^2 = \mathrm{Var}(V_3) = \frac{\alpha_3(\alpha_1 + \alpha_2)}{(\alpha_1 + \alpha_2 + \alpha_3)^2 + (\alpha_1 + \alpha_2 + \alpha_3 + 1)}.$$

It is easy to see why a Dirichlet distribution and its marginal distributions (the $\beta$'s) are so useful for modeling data related to proportions, thus increasing the possibility that given an election and the associated estimates made by one of the candidates, one member of the family of Dirichlet's can be found which models the distributions of proportions of the vote for the candidates in a fairly accurate manner. To illustrate the "maleability" of the Dirichlet distributions, consider the six distributions in Figure 5.3, where the values of the parameters $\alpha_1$, $\alpha_2$, and $\alpha_3$ which determine the distribution are listed above each graph. It should be obvious that a fairly large assortment of two-dimensional empirical distributions can be "fit" by the members of this family.[19]

---

[19] It is important to note that the model will not be rendered impractical if the candidate is ignorant of the nature of joint probability distributions. In fact, in order for this structure to be useful, it will suffice that the candidate have a reasonable understanding of the relationship between ballots in $S$ and the ($\beta$) marginal distributions of the Dirichlet distribution. Furthermore, his view of this relationship may be rather rudimentary. For example, candidate's conjectures such as, "If I choose this platform and my opponent chooses that one, then (1) the probability that my proportion of the vote is greater than forty percent is 0.8, (2) the probability that my opponent's proportion of the vote is greater than forty percent is 0.6, and (3) the probability that less than thirty percent of the electorate abstain is 0.6," will provide sufficient information to construct the Dirichlet distribution and compute the candidate's payoff with respect to a specified objective function. Furthermore, even though we indicate that the candidate chooses parameters in the space $A$, if a spatial formulation is employed, that is not technically correct. In that case, the candidate might think only in terms of ballots and payoffs, with his "resident theorist" keeping track of the mapping $\alpha_j$ from $S$ onto $A$. Hence, our model is more than a mathematical generalization of the candidate's decision process; if, in a given election, a spatial formulation is at all realistic, our model could be employed to assist a candidate in locating spatial equilibria.

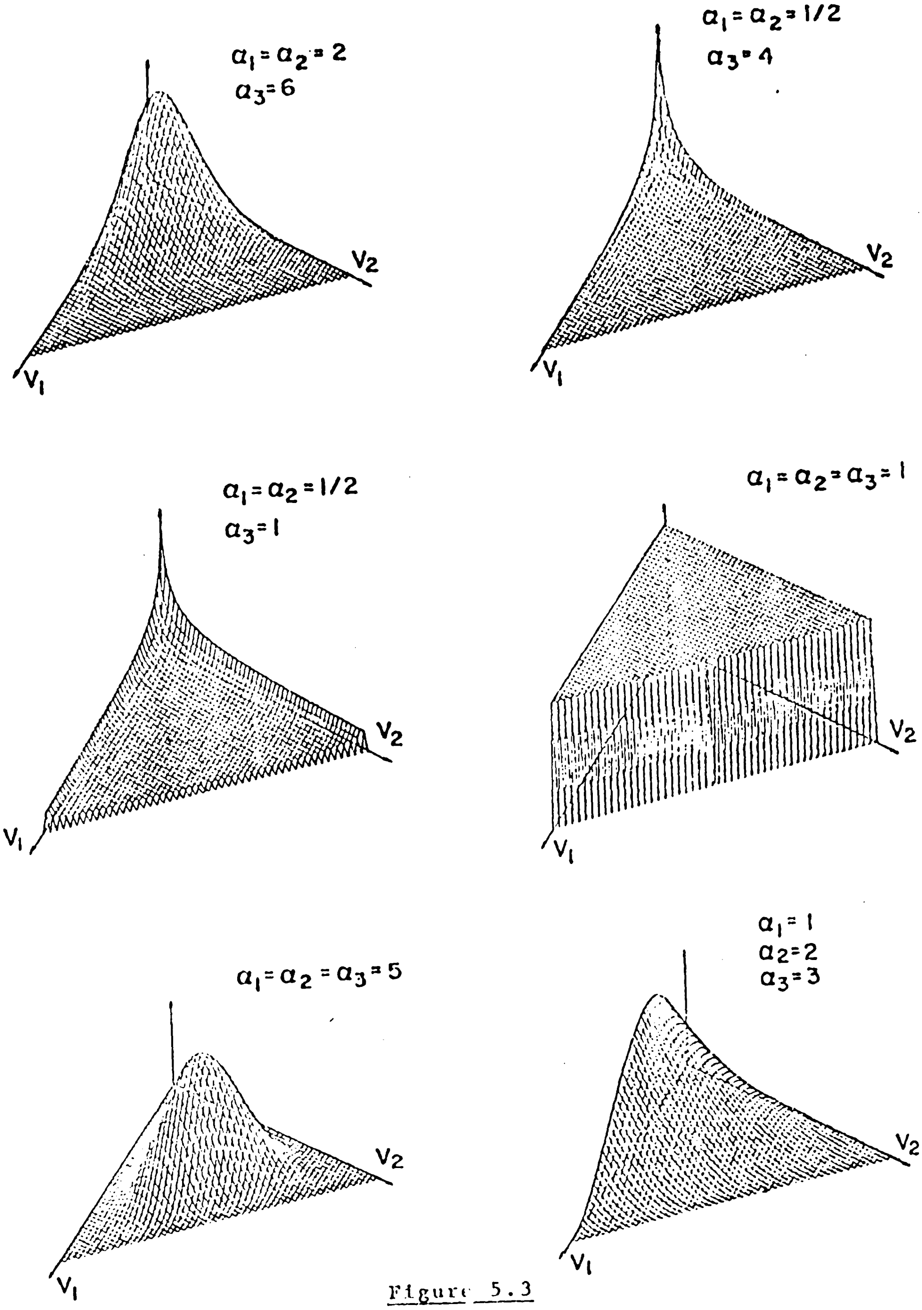



Figure 5.3

Examples of Dirichlet Distributions

We will need the following three lemmas in order to prove the principal results of this paper:

Lemma 5.1: If $(V_1, V_2)$ has a Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \alpha_3)$, then $V_1/(V_1 + V_2)$ has a beta distribution with parameters $(\alpha_1, \alpha_2)$.

Proof: If $(V_1, V_2)$ is Dirichlet with parameters $(\alpha_1, \alpha_2, \alpha_3)$, then there exist independent chi-square random variables $X_1$, $X_2$, and $X_3$ with degrees of freedom $\nu_1 = 2\alpha_1$, $\nu_2 = 2\alpha_2$ and $\nu_3 = 2\alpha_3$ respectively, such that $V_1 = X_1/(X_1 + X_2 + X_3)$ and $V_2 = X_2/(X_1 + X_2 + X_3)$. Since $V_1/(V_1 + V_2) = X_1/(X_1 + X_2)$ and since the beta distribution with parameters $(\alpha_1, \alpha_2)$ can be characterized as the distribution of $Y = X_1/(X_1 + X_2)$ where $X_1$ and $X_2$ are independent chi-square random variables with degrees of freedom equal to $\nu_1 = 2\alpha_1$ and $\nu_2 = 2\alpha_2$, the result follows.

Q. E. D.

Lemma 5.2: If $V$ has a beta distribution with parameters $(\alpha, \alpha')$, then $F_V(\lambda)$ is a class of decreasing functions in $\alpha$ for any fixed value of $\alpha'$.[20]

Proof: Suppose that $V \sim \beta(\alpha, \alpha')$.

Then there exist independent chi-square random variables

[20] $F_{V_1}(\lambda)$ is simply the cumulative distribution function of the random variable $V_1$ (which depends on $\alpha_1$). This lemma specifies that for any value of $\lambda$ and for fixed $\alpha'$, the sequence of functions $F_{V_1}(\lambda)$ is monotone increasing in $\alpha_1$.

$X_1$ and $X_2$ with degrees of freedom $\nu_1 = 2\alpha$ and $\nu_2 = 2\alpha'$ respectively, such that $V = X_1/(X_1 + X_2) = (1 + X_2/X_1)^{-1}$.
Therefore, $F_V(\lambda) = \Pr[v \leq \lambda] = \Pr[(1 + X_2/X_1)^{-1} \leq \lambda]$.
Let $\alpha'$ and $\lambda$ be fixed.
The degrees of freedom $\nu_1$ of $X_1$ increase without bound as $\alpha$ increases without bound. Furthermore, if $X_1 \sim \chi^2_{\nu_1}$, $X_1' \sim \chi^2_{\nu_1'}$, and $\nu_1 < \nu_1'$, then $\Pr[X_1 \leq \lambda] > \Pr[X_1' \leq \lambda]$ and $\Pr[1/X_1 \leq \lambda] < \Pr[1/X_1' \leq \lambda]$. Consequently, since $\alpha'$ (and, therefore, $\nu_2$) is fixed,
$\Pr[(1 + X_2/X_1)^{-1} \leq \lambda] > \Pr[(1 + X_2/X_1')^{-1} \leq \lambda]$.
This means that $F_V(\lambda)$ is a class of decreasing functions of $\alpha$.

Q. E. D.

<u>Lemma 5.3</u>: If $(V_1, V_2)$ has a Dirichlet distribution with parameters $(\alpha_1, \alpha_2, \alpha_3)$, then $F_{V_1-V_2}(\lambda)$ is a class of decreasing functions of $\alpha_1$ for any fixed values of $\alpha_2$ and $\alpha_3$.

<u>Proof</u>: As in the previous two lemmas, assume that $(V_1, V_2)$ is Dirichlet and let $X_1$, $X_2$, and $X_3$ be the appropriate independent chi-square distributions.
Then $V_1 - V_2 = (X_1 - X_2)/(X_1 + X_2 + X_3) = (1 + X_2/X_1)/(1 + (X_2 + X_3)/X_1)$.
Therefore,
$F_{V_1-V_2}(\lambda) = \Pr[v_1 - v_2 \leq \lambda] = \Pr[(1 + X_2/X_1)/(1 + (X_2+X_3)/X_1) \leq \lambda]$.
As $\alpha_1$ increases without bound, the degrees of freedom $\nu_1$ associated with $X_1$ also increase without bound.

Let $X_1 \sim \chi^2_{\nu_1}$ and $X_1' \sim \chi^2_{\nu_1'}$, and let $\alpha_2$, $\alpha_3$, and $\lambda$ be fixed.

Then $\alpha_1 < \alpha_1' \iff \nu_1 < \nu_1'$

$$\Rightarrow \Pr[X_1 \leq \lambda] > \Pr[X_1' \leq \lambda].$$

$$\Rightarrow \Pr[1+(X_2+X_3)/X_1 \leq \lambda] < \Pr[1+(X_2+X_3)/X_1' \leq \lambda].$$

$$\Rightarrow \Pr[(1-X_2/X_1)/(1+(X_2+X_3)/X_1) \leq \lambda] > \Pr[(1-X_2/X_1')/(1+(X_2+X_3)/X_1') \geq \lambda].$$

Consequently, $F_{V_1-V_2}(\lambda)$ is a class of decreasing functions of $\alpha_1$ for fixed $\alpha_2$ and $\alpha_3$.

Q. E. D.

The first theorem simply specifies that if the parameter space $A$ is unbounded, then equilibria do not exist. In general, equilibria, if they exist at all, will be found on the boundary of the parameter space; so if that space is not constrained, equilibrium strategies will not obtain.

Theorem 5.1: If the parameter space $A$ is unbounded and if the random pair $(V_1, V_2)$ has a Dirichlet distribution, then no (strong or weak) equilibrium exists for the objective functions $O_1$ through $O_7$.[21]

Proof: Without loss of generality, we will construct the proof from the perspective of the first candidate, $C_1$. We will assume that $A$ is unbounded, that $(V_1, V_2)$ is

---

[21]Recall that the $j^{th}$ candidate has a function $\alpha_j : S_j \to A_j$ where $S_j \subset S$ and $A_j \subset A$. In effect, when we specify that $A$ is unbounded, we actually mean that the range $A_j$ of $\alpha_j$ is unbounded for $C_j$, $j=1,2,\ldots,p$.

Dirichlet, and show that each objective function is an increasing function of $\alpha_1$.

For $O_1$ (Expected Plurality):

$$\phi_1(\underset{\sim}{\theta}) = E[V_1(\underset{\sim}{\theta}) - \max_{k\neq 1}\{V_k(\underset{\sim}{\theta})\}].$$

$$= \nu_1(\underset{\sim}{\theta}) - \nu_2(\underset{\sim}{\theta}).$$

$$= (\alpha_1 - \alpha_2)/(\alpha_1 + \alpha_2 + \alpha_3).$$

For any fixed values of $\alpha_2$ and $\alpha_3$, this is an increasing function of $\alpha_1$.

For $O_2$ (Proportion of the Expected Vote):

$$\phi_1(\underset{\sim}{\theta}) = E[V_1(\underset{\sim}{\theta})] \;/ \sum_{k=1}^{p} E[V_k(\underset{\sim}{\theta})].$$

$$= \alpha_1/(\alpha_1 + \alpha_2).$$

For any fixed value of $\alpha_2$, this is an increasing function of $\alpha_1$.

For $O_3$ (Expected Vote):

$$\phi_2(\underset{\sim}{\theta}) = E[V_1(\underset{\sim}{\theta})].$$

$$= \alpha_1/(\alpha_1 + \alpha_2 + \alpha_3).$$

For fixed values of $\alpha_2$ and $\alpha_3$, this is an increasing function of $\alpha_1$.

For $O_4$ (Probability that Plurality Exceeds Some Level):

$$\phi_1(\underset{\sim}{\theta}) = \Pr[V_1(\underset{\sim}{\theta}) - \max_{k\neq 1}\{V_k(\underset{\sim}{\theta})\} \geq \lambda_1].$$

$$= \Pr[V_1(\underset{\sim}{\theta}) - V_2(\underset{\sim}{\theta}) \geq \lambda_1].$$

$$= 1 - F_{V_1 - V_2}(\lambda_1).$$

By Lemma 5.3, for fixed values of $\alpha_2$, $\alpha_3$, and $\lambda_1$, this is an increasing function of $\alpha_1$.

For $O_5$ (Probability that Proportion Exceeds Some Level):

$$\phi_1(\underset{\sim}{\theta}) = \Pr[V_1(\underset{\sim}{\theta})/\sum_{k=1}^{p} V_k(\underset{\sim}{\theta}) \geq \lambda_1].$$

$$= \Pr[V_1(\underset{\sim}{\theta})/(V_1(\underset{\sim}{\theta}) + V_2(\underset{\sim}{\theta})) \geq \lambda_1].$$

By Lemma 5.1, $W = V_1/(V_2 + V_3) \sim \beta(\alpha_1,\alpha_2)$, so $\phi_1(\underset{\sim}{\theta}) = 1 - F_W(\lambda_1)$.

By Lemma 5.2, if $\alpha_2$ and $\alpha_1$ are fixed, $\phi_1(\underset{\sim}{\theta})$ is an increasing function of $\alpha_1$.

For $O_6$ (Probability that Vote Exceeds Some Level):

$$\phi_1(\theta) = \Pr[V_1(\underset{\sim}{\theta}) \geq \lambda_1].$$

But $V_1(\underset{\sim}{\theta})$ is simply a $\beta(\alpha_1, \alpha)$-distributed marginal distribution of the Dirichlet, $(V_1, V_2)$.

Therefore, $\phi_1(\underset{\sim}{\theta}) = 1 - F_{V_1}(\lambda_1)$, and, by Lemma 5.2, $\phi_1(\underset{\sim}{\theta})$ is an increasing function of $\alpha_1$ for any fixed value of $\alpha$ and $\lambda_1$.

For $O_7$ (Expected Proportion of the Vote):

$$\phi_1(\underset{\sim}{\theta}) = E[V_1(\underset{\sim}{\theta})/\sum_{k=1}^{p} V_k(\underset{\sim}{\theta})].$$

$$= E[V_1(\underset{\sim}{\theta})/(V_1(\underset{\sim}{\theta}) + V_2(\underset{\sim}{\theta}))].$$

$$= \alpha_1/(\alpha_1 + \alpha_2).$$

The last equality follows from Lemma 5.1,

$W = V_1/(V_1 + V_2) \sim \beta(\alpha_1, \alpha_2)$, and the fact that

$E[W] = \alpha_1/(\alpha_1 + \alpha_2)$.

Consequently, for any fixed value of $\alpha_2$, $\phi_1(\underset{\sim}{\theta})$ is an increasing function of $\alpha_1$.

We have shown that each of the seven objective functions is an increasing function of $\alpha_1$ for any fixed values of $\alpha_2$ and $\alpha_3$. Similarly, we can show that $\phi_2(\underset{\sim}{\theta})$ is an increasing function of $\alpha_2$ for any fixed values of $\alpha_1$ and $\alpha_3$. Therefore, if the parameter space $A$ is unbounded, no strong or weak equilibrium strategies exist.

Q. E. D.

It is noteworthy that in each of the cases in the proof of Theorem 5.1 $\lim_{\alpha_1 \to \infty} \phi_1(\underset{\sim}{\theta})$ exists. Despite this fact, in each instance $\phi_1(\underset{\sim}{\theta})$ is an increasing function of $\alpha_1$. The substantive interpretation of these two facts is that even though no equilibrium strategies exist for $C_j$--and he will always have an incentive to unilaterally relocate his spatial platform, essentially by taking successively larger values of $\alpha_j$--the marginal payoff for altering his strategy will become increasingly smaller. In fact, in this formulation it is possible, given any $\varepsilon > 0$, however small, to locate a non-equilibrium strategy such that the marginal payoff for <u>any</u> unilateral move by $C_j$ from that strategy will be less than $\varepsilon$. Consequently, if the model were extended to incorporate a cost to $C_j$ for relocating his platform, then there would be effective--if not formal--equilibria.

The next theorem shows that under appropriate conditions--including a bound on the parameter space $A$--equilibrium strategies will obtain.

<u>Theorem 5.2</u>: If $A$ is a convex, compact region, then there exist both strong and weak equilibria.[22]

Although the proof of this theorem is in the text of this dissertation, it will be instructive to illustrate it first with several examples. In Figure 5.4, the convex, compact parameter space is a hemisphere with its face parallel to the $\alpha_1\alpha_3$-plane. Starting at any point $(\alpha_{11}, \alpha_{21}, \alpha_{30})$ in the space, the first candidate $C_1$ will maximize his objective function, subject to $C_2$ choosing a strategy associated with $\alpha_{21}$, by "moving" to the weak equilibrium point $(\alpha_{12}, \alpha_{22}, \alpha_{30})$ on the surface of the hemisphere.[23] At $(\alpha_{12}, \alpha_{22}, \alpha_{30})$, $C_1$ cannot unilaterally increase his payoff; but $C_2$ can maximize his objective function, subject to $C_1$ choosing a strategy associated with $\alpha_{12}$ by "moving" to the weak equilibrium point $(\alpha_{13}, \alpha_{23}, \alpha_{30})$ on the face of the hemisphere. Again, however, $C_1$ will have an incentive to choose $(\alpha_1{*}, \alpha_2{*}, \alpha_{30})$ over $(\alpha_{13}, \alpha_{23}, \alpha_{30})$ since that will maximize his payoff, given $C_2$'s previous choice. Furthermore, since at $(\alpha_1{*}, \alpha_2{*}, \alpha_{30})$ neither candidate can increase his payoff by a unilateral "move" to a new platform,

---

[22] To be precise, we actually assume that the range $A_j$ of the function $\alpha_j$ is compact (closed and bounded) for each $j = 1,2,\ldots,p$.

[23] The point $(\alpha_{12}, \alpha_{22}, \alpha_{30})$ is a weak equilibrium for $C_1$, given that $C_2$'s choice is $\alpha_{22}$, because, as was shown in Theorem 5.1, each of the seven objective functions is maximized by choosing the largest attainable value of $\alpha_1$.

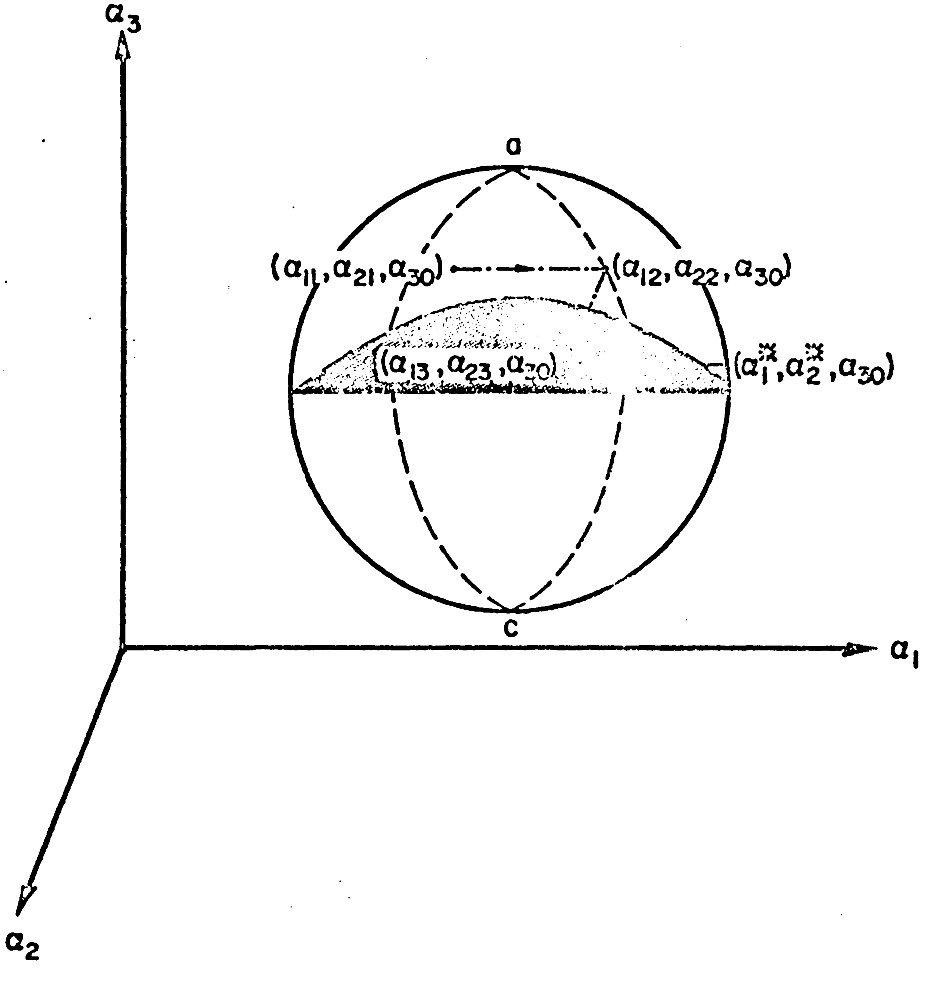



Figure 5.4

Equilibria in a Convex, Compact Region

it is a strong equilibrium. Finally, it is easy to see that in this example (1) the set $C$ of equilibria is the set of points on the semicircle $\widehat{abc}$; (2) given any value of $\alpha_3$, there is a unique equilibrium; (3) given any "starting point" $(\alpha_{11}, \alpha_{21}, \alpha_{30})$ there will be at most two intermediate weak equilibrium "moves" before a strong equilibrium will obtain; and (4) each equilibrium in $C$ can be obtained by choosing at least one "starting point" and employing the decision algorithm described above.[24]

Notice that in this example all of the action takes place in a cross-section of the parameter space defined by the intersection of $A$ with the plane through $\alpha_3 = \alpha_{30}$. In Figure 5.5, the triangle DEF is such a cross-section of a parameter space for a fixed value of $\alpha_3$, say $\alpha_{30}$. In this particular example, if we choose any point in the set except $E = (\alpha_1^*, \alpha_2^*)$ and employ the previously discussed algorithm,

---

[24] The example illustrated in Figure 5.4 is useful for demonstrating that after $C_2$ chooses a platform mapped by $\alpha_2$ into a point on the face of the hemisphere, he can do nothing else to increase his payoff. On the other hand, if $C_2$ can conduct a campaign which eventuates in either a very small or a very large turnout ($\alpha_3$ is either small or large), then he can effectively reduce $C_1$'s payoff ($\alpha_1$ will be small) by essentially forcing an equilibrium close to either a or c with no reduction in his own payoff. In short, the example illustrates what every candidate knows; that it is frequently useful to consider the effect of a large or small turnout on one's chances of winning (however winning is defined).

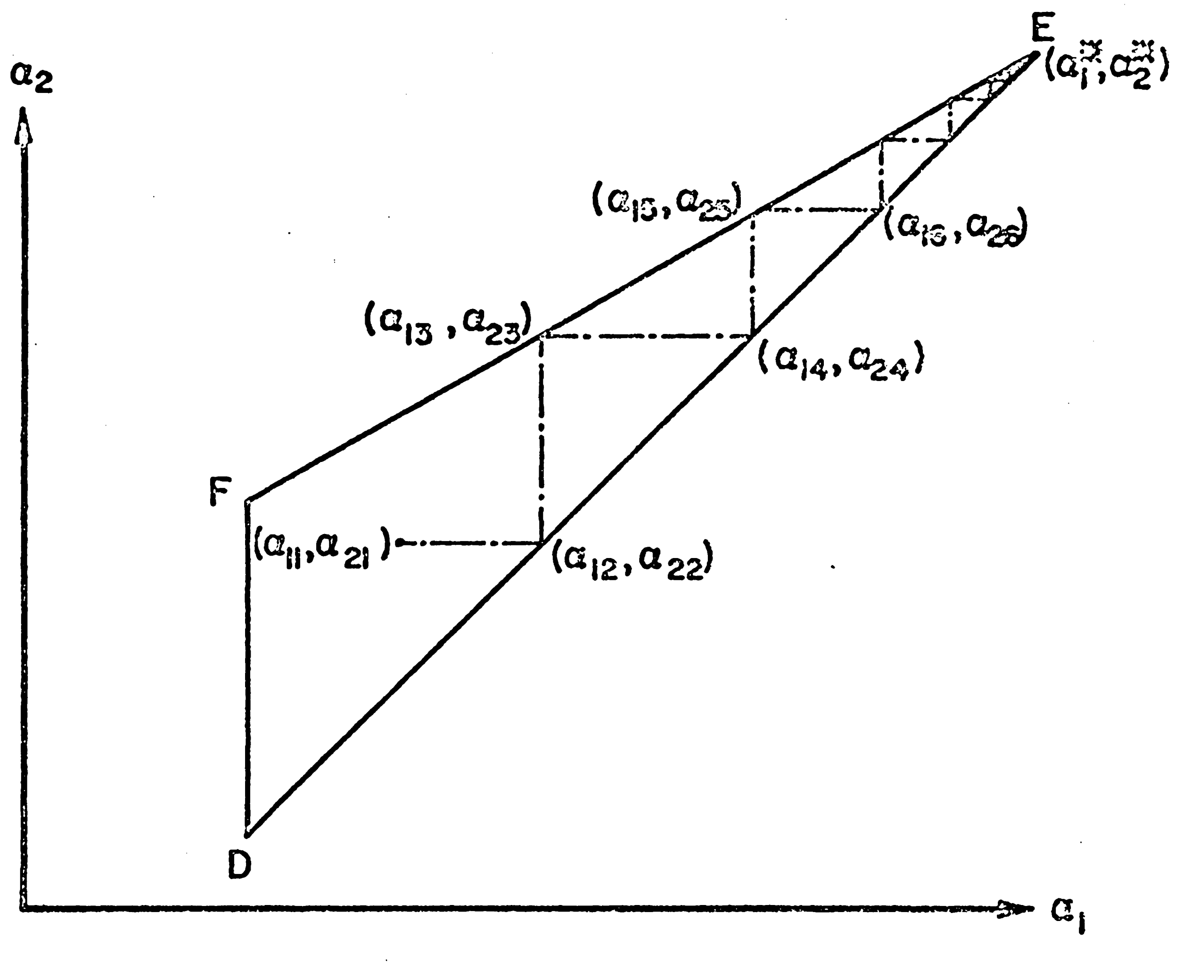



Figure 5.5

A Unique Equilibrium in the Cross-Section of a Convex, Compact Region we will not reach a strong equilibrium in a finite number of steps. On the other hand, since $A$ is a convex, compact (closed and bounded) space, the cross-section will also be a convex, compact set, and the monotone sequence $(\alpha_{1n}, \alpha_{2n})$ will converge to a point $(\alpha_1*, \alpha_2*)$.[25] It is easy to see that the point to which the sequence converges is an equilibrium point for any objective function which is maximized by increasing $\alpha_1$ (or $\alpha_2$), and it is also intuitively obvious that the equilibrium is on the boundary of the cross-section (and, consequently, on the boundary B of $A$). In this example, $E = (\alpha_1*, \alpha_2*)$ is the unique equilibrium for the specified value of $\alpha_3 = \alpha_{30}$. It is clear that if the candidates are at any point in the set except E, one or the other will have a unilateral incentive to choose another strategy; however, if they are located at E, neither will have such an incentive to "move" to another point in the space.

It is coincidental that in both of these examples there are unique equilibria for fixed values of $\alpha_3$, the parameter related to abstentions. In fact, if we take the parameter space to be the hemisphere in Figure 5.4 rotated ninety degrees about the $\overline{ac}$ axis, clockwise in the $\alpha_1\alpha_2$-plane, then

---

[25]The sequence $(\alpha_{1n}, \alpha_{2n})$ is monotone in the sense that $(\alpha_1, \alpha_2) \preceq (\alpha_1', \alpha_2')$ if and only if $\alpha_1 \le \alpha_1'$ and $\alpha_2 \le \alpha_2'$.

each cross-section will contain an infinite set of equilibria.

A formal proof of Theorem 5.2 follows:

Proof: Since we are concerned with a two-candidate election, $A \subset E^3$. If $(\alpha_1, \alpha_2, \alpha_3) \varepsilon A$, then $C_1$ has the power to manipulate $\alpha_1$, $C_2$ can manipulate $\alpha_2$, and $\alpha_3$ is related to the proportion of the electorate which abstains.

Let $(\alpha_{11}, \alpha_{21}, \alpha_{30}) \varepsilon A$, assume that $\alpha_{30}$ is fixed, and define rays parallel to the $\alpha_1$- and $\alpha_2$- axes respectively by

$$X_{\alpha_{10}\alpha_{20}} = \{(\alpha_1, \alpha_2, \alpha_{30}) \mid \alpha_2 = \alpha_{20} \text{ and } \alpha_1 > \alpha_{10}\}$$

$$Y_{\alpha_{10}\alpha_{20}} = \{(\alpha_1, \alpha_2, \alpha_{30}) \mid \alpha_1 = \alpha_{10} \text{ and } \alpha_2 > \alpha_{20}\}.$$

Let B be the boundary of $A$.

Starting at the point $z_1 = (\alpha_{11}, \alpha_{21}, \alpha_{30}) \varepsilon A$ we define a sequence of ordered triples as follows:

$z_2 = (\alpha_{12}, \alpha_{22}, \alpha_{30})$, where

$$\alpha_{12} = \max\{\alpha_1 \mid (\alpha_1, \alpha_{22}, \alpha_{30}) \varepsilon X_{\alpha_{11}\alpha_{21}} \cap B\} \text{ and } \alpha_{22} = \alpha_{21}.$$

$z_3 = (\alpha_{13}, \alpha_{23}, \alpha_{30})$, where

$$\alpha_{23} = \max\{\alpha_2 \mid (\alpha_{13}, \alpha_2, \alpha_{30}) \varepsilon Y_{\alpha_{12}\alpha_{22}} \cap B\} \text{ and } \alpha_{13} = \alpha_{12}.$$

$z_4 = (\alpha_{14}, \alpha_{24}, \alpha_{30})$, where

$$\alpha_{14} = \max\{\alpha_1 \mid (\alpha_1, \alpha_{24}, \alpha_{30}) \varepsilon X_{\alpha_{13}\alpha_{23}} \cap B\} \text{ and } \alpha_{24} = \alpha_{23}.$$

In general, $z_n = (\alpha_{1n}, \alpha_{2n}, \alpha_{30})$, where

$$\begin{cases} \alpha_{1n} = \max\{\alpha_1 \mid (\alpha_1, \alpha_{2n}, \alpha_{30}) \varepsilon X_{\alpha_{1n-1}\alpha_{2n-1}} \cap B\} \text{ and } \alpha_{2n} = \alpha_{2n-1} & \text{if } n \text{ is even.} \\ \alpha_{2n} = \max\{\alpha_2 \mid (\alpha_{1n}, \alpha_2, \alpha_{30}) \varepsilon Y_{\alpha_{1n-n}\alpha_{2n-1}} \cap B\} \text{ and } \alpha_{1n} = \alpha_{1n-1} & \text{if } n \text{ is odd.} \end{cases}$$

Letting $z = (\alpha_1, \alpha_2, \alpha_{30})$ and $z' = (\alpha_1', \alpha_2', \alpha_{30})$ we define a partial order on $A$ by $z \preceq z'$ if and only if $\alpha_1 \leq \alpha_1'$ and $\alpha_2 \leq \alpha_2'$, and define the distance $d(z, z')$ between $z$ and $z'$ to be the usual Euclidean distance, that is

$$d(z, z') = \sqrt{(\alpha_1 - \alpha_1')^2 + (\alpha_2 - \alpha_2')^2}.$$

With this, we claim that $(z_n)$ is a Cauchy sequence which is monotone increasing in $\preceq$.

The monotonicity of $(z_n)$ is true by construction. If $(z_n)$ were not Cauchy, then there exists an $\varepsilon > 0$ such that for any integer $m$ there exists $n > m$ such that $d(z_n, z_{n+1}) > \varepsilon$. But, by virtue of the Archimedian Principle, this means that given any arbitrary distance $\Delta$ from the origin vector $0$, there exists an integer $m$ such that for all $n > m$, $d(0, z_n) > \Delta$.

Since $A$ is bounded, this is a contradiction. Therefore, $(z_n)$ is a Cauchy sequence.

Now observe that $(z_n)$ is a monotone increasing, convergent sequence with each component (except possibly $z_1$) on the boundary $B$ of $A$. Since $B$ is a closed subset of $A$, the limit point $z^* = (\alpha_1^*, \alpha_2^*, \alpha_{30})$ of $(z_n)$ is also on $B$. Obviously, each component of the sequence $(z_n)$ (except possibly $z_1$) is a weak equilibrium, and we claim that $z^*$ is a strong equilibrium. If not, there exists $z^{**} = (\alpha_1^{**}, \alpha_2^{**}, \alpha_{30}) \varepsilon A$ such that $\alpha_1^{**} > \alpha_1^*$ and $\alpha_2^{**} = \alpha_2^*$ or else $\alpha_1^{**} = \alpha_1^*$ and $\alpha_2^{**} > \alpha_2^*$, i.e., one of the candidates has a unilateral incentive to move. It follows that $z^{**} \preceq z^*$. But since $z^* \varepsilon B$, the boundary of $A$, and since $A$ is a convex set, this means $z^{**} \notin A$, a contradiction. Consequently, $z^*$ is a strong equilibrium.

Q. E. D.

Finally, we will show that with fairly natural and realistic restrictions on the parameter space, all of the objective functions listed in Table 5.1 are equivalent for the $j^{th}$ candidate. In particular, we have

Theorem 5.3: In a two-candidate election, if the candidates' proportions of the vote have a Dirichlet distribution, and if the parameter space is a convex, compact region, then the objective functions $O_1$ through $O_7$ are equivalent.

In the proof of Theorem 5.1 it was shown that the $j^{th}$ candidate's payoff under each of the seven objective functions can be increased by increasing $\alpha_j$. In Theorem 5.2 it was demonstrated that if the parameter space is convex and compact, then each of the objective junctions is maximized by points $(\alpha_1^*, \alpha_2^*, \alpha_3)$ on the boundary of the parameter space. Consequently, the ballots $\underset{\sim}{\theta} \varepsilon S$ which are mapped into $(\alpha_1^*, \alpha_2^*, \alpha_3)$ are equilibrium strategies; and since the sets of equilibria are identical for each objective function, $O_1$ through $O_7$ are equivalent.

Q. E. D.

# CHAPTER VI

## SUMMARY AND CONCLUSIONS

In this dissertation we have attempted to answer three of the important questions regarding formal theories of voting under majority rule: (1) is it possible to develop a mathematically tractable spatial model of political participation in which citizens are heterogeneous both with regard to their preferences and the manner in which they "measure" the relative salience of the election issues; (2) is it possible to determine optimal candidate strategies in a spatial model (even if the citizens' perceptions of the relative importance of the issues is homogeneous) when citizens might abstain from voting because they are alienated from all competing candidates' platforms; and (3) are the sets of equilibrium ballots of candidates' strategies invariant over elections in which candidates attempt to maximize different objective functions?

Our answer to the first question is affirmative, and, thus, in that case we make a positive contribution to the mathematical theory of political behavior. For the model discussed in Chapter III we have shown that the generalized total median is the ultimate in strategic political platforms in a model where the (generic) $i^{th}$ voter may be characterized by the ordered pair $(x_i, A_i)$, where $x_i$ is his most preferred social state and $A_i$ is the $n \times n$ positive definite matrix by which the voter discriminates between the relative salience of the spatial issues. We argue that the importance of the mean preference has previously been exaggerated by confounding the analysis with both the assumption of common quadratic loss functions and symmetric distributions of voters' preferences. Under the assumption that political contests are between two candidates, several platforms are defined in Chapter III and the relationship between these platforms is investigated in a series of theorems and examples. Two displays summarize these relationships (Figures 3.6 and 3.9). It is noteworthy that many of the characteristics of a more rudimentary model developed by Hinich and Davis no longer hold in the more abstract spatial model.

Of course there are numerous restrictive assumptions supporting the model formulated in Chapter III, the most bothersome of which is the supposition that all citizens vote. It would be desirable to construct a model of electoral processes which incorporates individual differences regarding both preference specification and loss (or utility) specification and in which optimal strategies for candidates can be determined even when some citizens who are alienated from the vying candidates abstain from voting. We have been unable to construct such a model, but under the assumption that citizens have identical loss matrices, and, consequently, perceive the relative salience of election issues in an identical manner, Hinich and Ordeshook have attempted to formulate such a model. We argue that they have at least implicitly used two structurally different probability measures--a personal probability and a relative frequency probability--interchangeably while they are, in fact, not interchangeable. In addition, we have demonstrated that had they been consistent in utilizing only one kind of measure, their analysis--in either case--would lead to a voting model admitting interpersonal comparisons of utility. This is a serious departure from the theoretical intent of the original Davis-Hinich model, especially since the impetus for studying either their model or our generalized model is a desire to produce a mathemathical structure devoid of interpersonal comparisons of utility yet consistent with reality. In a sense then, we fail to make a positive contribution to spatial theories insofar as the question of voter abstentions is concerned. On the other hand, the first Hinich-Ordeshook paper on this topic appeared in 1969, so it is conceivable that since that time formal theorists with an interest in such problems have assumed that the problem is "solved." Our analysis essentially reopens this particular area of study. Furthermore, by identifying the problems with Hinich and Ordeshook's work we may provide formal modelers with some insight into a constructive, workable theory.

Our answer to the third question is, surprisingly, an affirmative one. It is surprising simply by virtue of the fact that our basic prejudice is that candidates who are attempting to maximize different objective functions will probably adopt different political platforms in order to do so, while the results of our formal analysis lead us to the opposite conclusion. We believe that it would be difficult to develop a more mathematically sophisticated model of the process discussed in Chapter V than the one which we formulate, yet it is the case that our formal analysis is inconsistent with our intuitive conceptualization of the actual behavior of candidates for office. We resolve this difficulty by claiming that our intuition is probably not too far off base, and conclude that a mathematical formulation of electoral processes in which candidates attempt to maximize different objective functions is not likely to reveal the complex nature of the candidates' behavior. Although our model would be useful for a candidate who is attempting to make gross estimates of the relative advantages of various platforms, if the candidate understands the characteristics of an election well enough to believe that equilibria are not invariant with respect to objective functions, then his own analysis is probably too sophisticated to be enlightened by our, or any other, formal model of this process.

In concluding this dissertation we would be remiss if we failed to express our prejudice about the importance of formal models to political scientists and politicians. Insofar as the utility of the spatial model for a candidate for office is concerned, we are not optimistic. The assumptions that (1) citizens have perfect information regarding the location of candidates' platforms, (2) citizens' preferences can be characterized by points in Euclidean n-space, and (3) citizens vote even when they are alienated from every candidate in a particular race are very restrictive. We are not alarmed about the assumption that voters measure the relative salience of the election issues with quadratic functions, because we believe that any model which provides for more general input data (for example, by allowing voters to have arbitrary convex loss functions) may be too complex to be verified empirically. It is not clear that, even with the aid of computers, there is any way to report, much less aggregate them. At any rate, it seems likely that the complexity of the input coupled with at least two unrealistic assumptions limits the usefulness of the spatial model for a candidate for office.

From another perspective, however, formal models in general, and particularly the spatial model, give political scientists insight into the structure of various political phenomena that they are unlikely to obtain by philosophizing in a less formal manner about those phenomena or by analyzing data relative to them.

We feel that the model explicated in Chapter V--which is based upon the Dirichlet distribution--may have some practical utility for politicians. More specifically, all that is required to obtain "optimal" platforms utilizing our model is that the candidate be reasonably adept at estimating the proportions of the vote which he and his opponents will receive and the proportion of the electorate which will abstain from voting if certain ballots are "chosen" by the candidates. While such computations are non-trivial, it is altogether possible that a thoughtful candidate working in consort with someone who understands the nature of the model could order the desirability of various platforms. Of course, in practice only platforms which are quite different could be ranked in this manner; nevertheless, we imagine that utilization of the model may assist the candidate under such conditions.

## REFERENCES


[ 1] Aranson, Peter; Hinich, Melvin; and Ordeshook, Peter. "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives." American Political Science Review, 68 (March, 1974).

[ 2] Arrow, Kenneth. Social Choice and Individual Values. 2nd ed. New York: John Wiley and Sons, 1963.

[ 3] Black, Duncan. The Theory of Committees and Elections. Cambridge: Cambridge University Press, 1958.

[ 4] Buchanan, James, and Tullock, Gordon. The Calculus of Consent. Ann Arbor: The University of Michigan Press, 1967.

[ 5] Davis, Otto; DeGroot, Morris; and Hinich, Melvin. "Social Preference Orderings and Majority Rule." Econometrica, 40 (January, 1972).

[ 6] Davis, Otto, and Hinich, Melvin. "A Mathematical Model of Policy Formation in a Democratic Society." Mathematical Applications in Political Science II. Edited by Joseph Bernd. Dallas: Southern Methodist University Press, 1966.

[ 7] Davis, Otto, and Hinich, Melvin. "On the Power and Importance of the Mean Preference in a Mathematical Model of Democratic Choice." Public Choice, 5 (Fall, 1968).

[ 8] Davis, Otto; Hinich, Melvin; and Ordeshook, Peter. "An Expository Development of a Mathematical Model of the Electoral Process." American Political Science Review, 64 (1970).

[ 9] Downs, Anthony. Economic Theory of Democracy. New York: Harper and Row Publishers, 1957.

[10] Farquharson, Robin. Theory of Voting. New Haven: Yale University Press, 1969.

[11] Good, I. J., and Mayer, Lawrence. "On Surfaces of Constant Societal Loss in a Model of Social Choice." Journal of Mathematical Sociology, 2 (August, 1972).

[12] Good, I. J., and Mayer, Lawrence. "Theorems on Condorcet Sets." Unpublished manuscript, 1972.

[13] Hinich, Melvin; Ledyard, John; and Ordeshook, Peter. "A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games." Journal of Politics, 35 (February, 1973).

[14] Hinich, Melvin, and Ordeshook, Peter. "Abstentions and Equilibrium in the Electoral Process." Public Choice, 7 (Fall, 1969).

[15] Hinich, Melvin, and Ordeshook, Peter. "Plurality Maximization Vs. Vote Maximization: A Spatial Analysis with Variable Participation." American Political Science Review, 64 (September, 1970).

[16] Hoyer, R. W., and Mayer, Lawrence. "Comparing Strategies in a Spatial Model of Electoral Equilibrium." American Journal of Political Science, 18 (August, 1974).

[17] Hoyer, R. W., and Mayer, Lawrence. "Further Results on Condorcet Sets." Unpublished manuscript, 1973.

[18] Hoyer, R. W., and Mayer, Lawrence. "On Social Preference Orderings under Majority Rule." Econometrica, 43 (July, 1975).

[19] Johnson, Norman, and Kotz, Samuel. Distributions in Statistics: Continuous Multivariate Distributions. New York: John Wiley and Sons, 1972.

[20] Johnson, Norman, and Kotz, Samuel. Distributions in Statistics: Continuous Univariate Distributions - 2. New York: John Wiley and Sons, 1970.

[21] Kramer, Gerald. "Sophisticated Voting over Multidimensional Choice Spaces." Journal of Mathematical Sociology, 3 (1973).

[22] Luce, R. Duncan, and Raiffa, Howard. Games and Decisions. New York: John Wiley and Sons, 1957.

[23] Miller, Philip. "Criticisms of McKelvey's Theory of Policy Related Voting." Mimeo., 1976.

[24] McKelvey, Richard. "Policy Related Voting and Electoral Equilibrium." Econometrica, 43 (1975).

[25] Pattanaik, Prasanta. Voting and Collective Choice. Cambridge: Cambridge University Press, 1971.

[26] Plott, Charles. "A Notion of Equilibrium and Its Possibility under Majority Rule." American Economic Review, 57 (1967).

[27] Quirk, James, and Saposnik, Rubin. Introduction to General Equilibrium Theory and Welfare Economics. New York: McGraw-Hill Book Co., 1968.

[28] Sen, Amartya. Collective Choice and Social Welfare. San Francisco: Holden-Day Inc., 1970.

[29] Shepsle, Kenneth. "Theories of Collective Choice." Political Science Annual, V: Collective Decision-Making. Edited by Cornelius Cotter, et al. Indianapolis: Bobbs-Merrill, 1973.

[30] Wendell, R. E., and Thorson, S. J. "Some Generalizations of Social Decisions under Majority Rule." Paper presented at the Annual Meeting of the Midwest Political Science Association, 1973.

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2

SOME MULTIVARIATE PROBLEMS

OF A SPATIAL MODEL OF VOTING

UNDER MAJORITY RULE

by

R.W. Hoyer

(ABSTRACT)

This dissertation is concerned with a number of outstanding problems in the analysis of collective decision-making. More specifically, we specify the mathematical foundation for a spatial model of voting under majority rule, and, building upon that foundation, we make three contributions to the formal theory of public choice.

First, we formulate a spatial model in which voters are heterogeneous both with respect to their most preferred social states and the manner in which they measure loss relative to their preferences, and within the structure of this model we identify various optimal strategies for the candidates for office.

Second, we show that all existing spatial models which purport to encompass voting abstentions by citizens who are alienated from the candidates are either mathematically inaccurate or else they include interpersonal comparisons of utility by the electorate.

Finally, we show that the set of optimal strategies for a candidate is invariant with respect to any one of seven objective functions which he may be attempting to maximize. This result is dependent upon the utilization of the family of Dirichlet distributions to model the joint distribution of the estimated proportions of the vote which the candidates will receive.