Voting Games with Incomplete Information

Thesis by

John W. Patty

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2001
(Submitted April 23, 2001)

Acknowledgements

I would like to thank the California Institute of Technology and the Alfred P. Sloan Foundation for their financial and logistical support of my work.

None of this research would have been possible without the unfaltering support of my advisor, Tom Palfrey. Richard McKelvey provided hours of stimulating discussion and insightful comments as well as co-authoring Chapter 3 with me. Jeff Banks provided invaluable advice and comments on the body of this work. In addition, R. Michael Alvarez, Fred Boehmke, Kim Border, Daniel Diermeier, John Duggan, Mark Fey, Sean Gailmard, Garrett Glasgow, Serena Guarnaschelli, Angela Hung, John Ledyard, Ken Schotts, Ken Shepsle, Bob Sherman, and Rob Weber all provided helpful comments and criticism on various parts of this research. Special thanks are due to John Duggan for finding errors in two proofs in an earlier draft of Chapter 2. The advice and support of Rhett Brown, John Miller, Alison Monahan, Nienke Oomes, Scott Page, Elizabeth Penn, and my family have been instrumental in the completion of not only this work, but much more.

This work is dedicated to the memory of Jeff Banks, whose scholarship, integrity, and life continue to inspire so many.

Abstract

We examine voting situations in which individuals have incomplete information over each others' true preferences. In many respects, this work is motivated by a desire to provide a more complete understanding of so-called probabilistic voting.

Chapter 2 examines the similarities and differences between the incentives faced by politicians who seek to maximize expected vote share, expected plurality, or probability of victory in single member, single vote, simple plurality electoral systems. We find that, in general, the candidates' optimal policies in such an electoral system vary greatly depending on their objective function. We provide several examples, as well as a genericity result which states that almost all such electoral systems (with respect to the distributions of voter behavior) will exhibit different incentives for candidates who seek to maximize expected vote share and those who seek to maximize probability of victory.

In Chapter 3, we adopt a random utility maximizing framework in which individuals' preferences are subject to action-specific exogenous shocks. We show that Nash equilibria exist in voting games possessing such an information structure and in which voters and candidates are each aware that every voter's preferences are subject to such shocks. A special case of our framework is that in which voters are playing a *Quantal Response Equilibrium* (McKelvey and Palfrey (1995), (1998)). We then examine candidate competition in such games and show that, for sufficiently large electorates, regardless of the dimensionality of the policy space or the number of candidates, there exists a strict equilibrium at the social welfare optimum (i.e., the point which maximizes the sum of voters' utility functions). In two candidate contests we find that this equilibrium is unique.

Finally, in Chapter 4, we attempt the first steps towards a theory of equilibrium in games possessing both continuous action spaces and action-specific preference shocks. Our notion of equilibrium, *Variational Response Equilibrium*, is shown to exist in all

games with continuous payoff functions. We discuss the similarities and differences between this notion of equilibrium and the notion of Quantal Response Equilibrium and offer possible extensions of our framework.

Contents

[1] I am extremely grateful to John Duggan for many helpful comments on this topic, including the discovery of a serious error in an earlier draft of this section.

[2] **This chapter is jointly authored with Richard McKelvey, who acknowledges the financial support of NSF grant #SBR-9631627 to the California Institute of Technology.**

Chapter 1 Introduction

Games of incomplete information arise in many settings. Perhaps the most pervasive of such settings is environments in which many individuals' actions each affect one another. Such environments are inherently characterized by incomplete information since it is usually impossible for any individual to know every other individual's preferences or motives, for example.

In the following chapters, we examine an application of the theory of games of incomplete information to voting situations. In particular, we examine models in which proposals are made to a group of voters whose true preferences over outcomes are incomplete information.

We attempt to answer several questions, including examining the incentives to candidates in elections where candidates have incomplete information about voters' true preferences (Chapter 2), characterizing a Nash equilibrium of spatial voting games when the number of voters is large (Chapter 3), and defining a notion of equilibrium for games of incomplete information in which there exists a metric on the action space (Chapter 4).

1.1 Literature Review

This work is closely related to that of many other scholars. This section attempts to review previous work. The relationships between this work and earlier results are outlined when helpful.

1.1.1 Related Work on Games of Incomplete Information

Most of this research is firmly embedded in a Bayesian environment. As such, this work would not have been possible without the pioneering research of Harsanyi (1967-68), which defined the primitives and earliest solution concepts for games of incom-

plete information.

Similarly, most of the primitives used in this research satisfy the conditions imposed by Milgrom and Weber (1985) in a seminal work on distributional strategies and Bayesian Nash equilibria in games of incomplete information. We use their existence and purification theorems in Chapter 3 and provide results similar to theirs in Chapter 4.

1.1.2 Related Work on Candidate Objectives

Several articles discussing properties of different candidate strategies were published in the 1970s. Foremost among these early efforts is Aranson, Hinich, and Ordeshook (1974). Aranson, *et al.* offer an equivalence result which rests on assumptions regarding perturbations of the candidate's objective functions, perhaps representing forecast errors. Their result, however, requires that these forecast errors are unbiased and, more importantly, that the errors are uncorrelated with the strategies chosen by the candidates. As the authors point out, this assumption is untenable, since the value of the objective functions (even after the errors are taken into account) must fall between zero and one. A second equivalence result obtained by Aranson, *et al.* requires that the votes received in a two candidate election be distributed according to a multivariate normal distribution. This obviously requires that negative vote totals be a positive probability event. Aranson, *et al.* were unable to offer any equivalence results between expected plurality and probability of victory based on assumptions regarding the primitives of the model.

Hinich (1977), however, provided justification for examining expected vote share in place of probability of victory which depended only on the Central Limit Theorem. Hinich's equivalence result states that the two objective functions converged in 2 candidate elections without abstention. This finding was extended by Ledyard (1984) to include 2 candidate elections in which abstention is allowed. Providing intuition for his claim, Ledyard argues at the limit, which is never actually realized in his framework. In addition, there is a discontinuity at the limit, making his argument impossible to generalize immediately for finite numbers of voters.

1.1.3 Related Work on Probablisitic Voting

Much other literature has studied probabilistic voting (see Coughlin (1992) for a review of this literature). Hinich (1977) showed that the median voter theorem does not always hold in a setting with probabilistic voting, and he constructed examples in a one-dimensional space with equilibria at other locations. In particular, with quadratic utility functions, he obtained an equilibrium in two candidate elections at the mean (which is the social welfare optimum with those preferences). Coughlin and Nitzan (1981a), (1981b) (see also Coughlin (1992), p. 96, Theorem 4.2) proved if voters have likelihood of voting functions satisfying the Luce axioms over subsets, there is a local equilibrium at a point maximizing the social log likelihood. While this work was not explicitly rooted in a utility maximization framework, subsequent work (see (1992), p. 99-100, Corollaries 4.4 and 4.5, Theorem 4.2) shows how it can be so interpreted. Coughlin (1992) also gives various conditions on voter likelihood functions or on preferences that result in a global equilibrium. If the likelihood functions are concave, there is a global equilibrium. In a re-distributional model where voters have logarithmic utility functions for income, and candidates use a logistic model to estimate the probability that voters vote for each candidate, there is a global equilibrium at the social utility maximum (p. 57, Theorem 3.7). All of the above results are for two candidate competition. Recently, Lin, Enelow and Dorussen (1999) show that one can also obtain equilibrium for multi-candidate elections using probabilistic voting models. They assume preferences based on distance, with a random utility shock, and obtain local equilibria at the social utility maximum. Lin, et al. also find that if the utility shocks have high enough variance, then the expected vote function for each candidate becomes concave, implying the existence of a global equilibrium.

Recent work by Banks and Duggan (1999) examines the properties of spatial competition between two candidates in a very general class of models of probabilistic voting. They assume that voters' probabilistic behavior is a function of the preferences over the policy space, rather than the expected utility of each action. Thus, voters' behavior is invariant to the probability that their vote will have any effect. Banks and

Duggan show that the point which maximizes the sum of the voters' utility functions has a special property, in that it is, very generally, a critical point of both candidates' expected vote share. They also provide results concerning the continuity of "almost core" policies when the core "almost" exists, where almost is defined topologically on the space of possible policies.

In all of the above cited probabilistic voting literature, game theoretic considerations for the voter are not modeled. Voters are assumed to vote based on their preferences for the candidate policy positions rather than based on the effect their vote will have on the outcome of the election. Ledyard (1984) develops a Bayesian model of two candidate competition that does model the game theoretic considerations for the voter. In his model, voters vote deterministically (there is no random utility shock to preferences), but they can abstain as well as vote for one of the two candidates, and the cost of voting is a random variable. Voter types consist of preferences as well as a cost of voting. He shows that in large elections, if voting costs are non-negative, there is an equilibrium at the social welfare optimum, which under certain restrictive conditions on the distribution of costs, is a global equilibrium. Myerson (1997) extends Ledyard's results in a model where the number of voters is a Poisson random variable, unknown to the voters. He shows that as long as the density function of the costs of voting is positive at zero, there is a global equilibrium in Ledyard's model as the number of voters becomes large. The Ledyard model, as well as Myerson's generalization of it, require that no voters have negative costs of voting.

1.2 Notation

This section defines much of the notation used in the remainder of this work. c

1. Topology, Sets, etc.

   We denote the set of M-dimensional real vectors with all strictly positive entries by $\mathbb{R}^M_{++}$ and the set of positive integers by $\mathbb{Z}_{++}$.

Let X be a space and χ be a topology on X. In an abuse of notation, we denote the Borel σ-algebra of (X, χ) by $\mathcal{B}(X)$. For any set $B \in \mathcal{B}(X)$, the *indicator function* defined by B is denoted by $\mathbf{1}[B]$. For any finite set Z, we denote the set of probability distributions over Z by $\Delta(Z)$. If X is a metric space, we will write $||\cdot, \cdot||$ to denote the metric on X.

Given a pair of topological spaces X and Y, will write $\mathcal{C}(X, Y)$ to denote the set of continuous functions taking X as their domain and Y as their range.

2. Convergence

When discussing convergence, we use $\rightarrow$ to denote pointwise convergence of functions and $\rightarrow\rightarrow$ to denote weak convergence of measures. For any real-valued function g taking as its domain a subset of a finite dimensional Euclidean space, we adopt the notation $\lim g(x) = z$ if for any $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$, such that for any point y satisfying $||x - y|| < \delta(\varepsilon)$, $|g(x) - g(y)| < \varepsilon$.

3. Vectors and Matrices

For any vector $x \in \mathbb{R}^m$, we write $x = (x^1, \ldots, x^m)$. Similarly, for any matrix $Y \in \mathbb{R}^m \times \mathbb{R}^n$, we write y^i_j for the entry in the i^{th} row and j^{th} column.

4. Derivatives

For any space $X \in \mathbb{R}^m$, and any continuously differentiable function $f : X \rightarrow \mathbb{R}$, we denote the vector of partial derivatives of f with respect to x by

$$Df = \left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^m} \right),$$

and write $Df(z)$ when this vector is evaluated at a point $z \in X$. Similarly, we will write $D^2 f$ for the matrix of second order partial derivatives of f with respect to x, and $D^2 f(z)$ for the evaluation of this matrix at a point $z \in X$.

5. Integration

For any integrable function f and probability measure ϕ, each defined on a topological space X, we denote the integral of f with respect to ϕ on $Y \subset X$

by one of the following notations:

$$\int_Y f\,d\mu,$$

$$\int_Y f(x)\mu(dx),$$

or

$$E_\mu[f(x)|x \in Y].$$

When $Y = X$, we may drop the notation for the space over which integration is performed if the context is clear.

Chapter 2 Candidate Objective Functions

2.1 Introduction

Game theory was first applied to the social sciences in an economic setting (von Neumann and Morgenstern (1944)), so the question of which payoff functions to use was often simple. While the debate regarding what actual preferences over monetary rewards look like is still unresolved, at least the supposition that preferences are weakly increasing in monetary rewards seemed uncontroversial.

When political scientists started to apply game theoretic models to the study of elections, the issue of payoff functions became more complicated, however. In particular, what should we assume that candidates wish to maximize as a result of their platform choice? On the one hand, each vote carries equal weight *ex ante*, so it seems intuitive that a candidate may seek to maximize her vote share. In addition, there are other reasons, such as "mandate" effects, reputation benefits in repeated elections, and signaling value to receiving more votes in an election (or poll) which may justify seeking to maximize the absolute amount of electoral support.

On the other hand, the first works in formal political theory assumed that candidates were pure office seekers (see Downs (1957) and Black (1958), for example), implying that candidates should maximize the probability of winning the election in a one-shot electoral model. The question of equivalence is straight forward: which, if any, different candidate objectives lead to identical optimal behavior?

While the question of equivalence is interesting for several reasons, including some normative reasons, such as what types of candidate objective functions lead to more "representative" outcomes in equilibrium, there are several technical reasons which motivate our examination. First, much of the work on spatial competition (see, for example, Hotelling (1929)) was in an economic environment, where the analogue to candidates are firms seeking to maximize profits. Thus, to apply the results of this *work directly* to political science, one must assume that candidates wish to maximize the number of votes they each receive. Second, the calculation of the expected number of votes is a much simpler operation than the calculation of the probability of victory in many models of electoral competition. Finally, even if the calculation of probability of victory is not difficult, this function is discontinuous in many electoral models in which the candidates' expected vote shares and expected pluralities are continuous.

As noted above, both Hinich (1977) and Ledyard (1984) provide equivalence results for elections with two candidates. Part of this chapter's motivation is to examine their results at a deeper level. In particular, we examine the question of best response equivalence in general probabilistic voting models. While we find difficulties with respect to both Hinich's and Ledyard's results, neither Hinich nor Ledyard focused much attention on equivalence as their results did not depend on any type of equivalence holding. Hinich's proposition is merely a statement, as the paper's principal finding (a one-dimensional spatial model in which the unique equilibrium is not at the median, but rather at the mean of the voters' ideal points) is important regardless of what the candidates' objective functions are assumed to be. Ledyard's results also do not require an equivalence result to be considered important.

2.1.1 Overview and Structure

Section 2.2 sets up a general model of voter behavior within single vote, single member, simple plurality electoral systems. Three different candidate objectives, maximization of expected plurality, maximization of expected vote share, and maximization of probability of victory, are formally defined in Section 2.3. We then provide a preliminary result which states that expected vote and expected plurality maximization are equivalent in all two candidate elections in which abstention is not allowed. The following section, Section 2.4, then formally defines four types of equivalence: best response and equilibrium equivalence in both pure and mixed strategies by the candidates. We prove, for completeness, a simple nesting result (Theorem 6) which essentially states that best response equivalence implies equilibrium equivalence between any pair of objective functions.

Section 2.5 contains one of our main results: establishing sufficient conditions for best response equivalence to hold between all three objective functions (Theorem 13). It turns out that our sufficient conditions are very stringent: there may only be two candidates, and each voter must act independently and, in expectation, identically. We refer to any electoral system in which voter behavior satisfies our sufficient conditions as possessing a representative voter. The remainder of Section 2.5.1 provides examples to show why these conditions are, at least in some sense, tight.

In Section 2.5.3, we examine Ledyard's (1984) equivalence result and provide an example highlighting a difficulty with equivalence in 2 candidate elections in which abstention is allowed. In particular, the special role of abstention and its effect on the probability of victory by a front-running candidate are discussed.

In Section 2.5.4 we examine Hinich's (1977) equivalence result. We provide a counterexample to his claim which raises a difficulty with extending our sufficient conditions to his framework.

Following the question of best response equivalence, we examine the question of equilibrium equivalence in Section 2.6. We provide sufficient conditions for local equilibrium equivalence between maximization of expected vote share and probability of victory in a broad class of models of electoral competition (Theorems 28 and 29), furthering recent work by Duggan (2000) on the topic.

Section 2.6.5 contains a discussion of necessary conditions for best response and equilibrium equivalence. We utilize a notion of genericity for infinite dimensional spaces, *shyness*, due to Hunt, Sauer, and Yorke (1992), and recently generalized by Anderson and Zame (2000), to show that the set of electoral games in which the necessary first order conditions for either type of equivalence are satisfied is small in a formal, measure-theoretic sense (Theorems 37 and 38). Section 2.7 contains a discussion of possible generalizations and extensions of our results to other models of voter behavior and other models of electoral competition. Section 2.8 concludes.

2.2 The Model

In this section we define the framework in which we will examine the question of equivalence between objective functions.

2.2.1 Primitives

Let $\mathcal{N}$ denote a finite set of voters, with $|\mathcal{N}| = N$, and $\mathcal{J}$ denote a finite set of alternatives (which may include both candidates and abstention, for example), with $|\mathcal{J}| = J$. We denote the set of *candidates* by $\mathcal{J}^0 \subset \mathcal{J}$, with $|\mathcal{J}^0| = J^0$. Each candidate j chooses a *policy*, $x_j \in X$, where X, a compact subset of a complete and separable metric space, denotes the policy space. We denote the J^0-dimensional vector of all policies by x, and the space of all such vectors is denoted by $Y = X^{J^0}$. The vector of all announced policies, other than the policy announced by candidate j, is denoted by x_{-j}, and the space of all such vectors by Y_{-j}. We denote the set of all mixed strategies for candidate j by M_j, the set of all profiles of mixed strategies by $M = \prod_{j=1}^{J^0}$, and set of all profiles of mixed strategies for candidates other than j by M_{-j}.

Each voter i chooses one alternative, denoted by $a_i \in \mathcal{J}$. The vector of all choices, $(a_1, \ldots, a_N)$, is denoted by a. The space of all such vectors is denoted by A. Each candidate j possesses an *objective function* $u_j : A \to \mathbb{R}$. For any $a \in A$ and $j \in \mathcal{J}$, we denote the *vote total* of candidate j by $v_j = \sum_{i=1}^{N} \mathbf{1}[a_i = j]$.

Each candidate j picks a policy proposal $x_j \in X$ simultaneously. These choices are then made common knowledge to the voters. After observing the policy proposals, each voter votes for one alternative. As described above, this vector of choices is denoted by a. For any $a \in A$, let $w(a) \in \{j \in \mathcal{J} | v_j \geq \max_{l \in \mathcal{J}} v_l\}$ denote the winning candidate at s. In the case of a tie, the winner is assumed to be determined by a fair lottery between all candidates j for which $v_j = \max_{l \in \mathcal{J}} v_l$. We denote the set of such candidates by $W(a)$. That is, we are examining a single winner, simple plurality rule system.

2.2.2 Voter Behavior

Each voter $i \in \mathcal{N}$ is characterized by a *response function*, $p_i : Y \to \Delta(\mathcal{J})$. Such a function represents the strategy of voter i. We denote the probability an alternative $j \in \mathcal{J}$ receives voter i's vote, conditional on policy proposal vector x, by $p_i^j(x)$.

Throughout, we assume that each $p_i(x)$ characterizes an independent multinomial random variable $a_i(x)$. This is stated formally below.

Assumption 1 (Independence) *Conditional on a vector of policy proposals, $x \in Y$, the set of $a_i(x)$ are mutually independent random variables, each distributed according to $p_i(x)$, respectively, for all $i \in \mathcal{N}$.*

Finally, we define an electoral game to be any sextuple, $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$. We denote the set of vectors of pure strategy Nash equilibria to an electoral game Γ by $PNE(\Gamma) \subset Y$ and the set of all Nash equilibria by $NE(\Gamma) \subset M$.

2.3 Candidate Objective Functions

We now use the set of p_i to define three different candidate objective functions, expected vote share, expected plurality, and probability of victory. For clarity, we first define these objective functions with respect to pure strategies by the candidates and then extend these definitions to the case of mixed strategies by the candidates.

2.3.1 Maximizing Expected Vote Share

Given opponents' pure strategies x_{-j}, an *expected vote share maximizing candidate* $j \in \mathcal{J}^0$ seeks to maximize

$$V_j(x_j; x_{-j}) = \frac{1}{N} E\left[\sum_{i=1}^{N} \mathbf{1}\left[a_i = j\right]\right],$$

which, given Assumption 1, reduces to

$$V_j(x) = \frac{1}{N}\sum_{i=1}^{N} p_i^j(x). \tag{2.1}$$

A few notes are in order. A candidate who seeks to maximize vote share is, in some sense, competing against all other alternatives – not just the other candidates. In particular, a candidate who seeks to maximize her expected vote share is also trying to increase the number of voters who turn out to vote for her. In addition, an expected vote share maximizing candidate is indifferent with respect to the distribution of the votes she does not receive.

2.3.2 Maximizing Expected Plurality

Given pure strategies x_{-j} by her opponents, an *expected plurality maximizing candidate* $j \in \mathcal{J}^0$ seeks to maximize

$$\pi_j(x_j; x_{-j}) = \frac{1}{N}E\left[\sum_{i=1}^{N}\mathbf{1}\{a_i = j\} - \max_{k\in\mathcal{J}^0, k\neq j}\left[\sum_{i=1}^{N}\mathbf{1}\{a_i = k\}\right]\right],$$

which, given Assumption 1, reduces to

$$\pi_j(x) = \frac{1}{N}\left(\sum_{i=1}^{N} p_i^j(x) - E\left[\max_{k\in\mathcal{J}^0, k\neq j}\left[\sum_{i=1}^{N}\mathbf{1}\{a_i = k\}\right]\right]\right). \tag{2.2}$$

That is, an expected plurality maximizing candidate seeks to maximize the difference between her own vote share and the maximum vote share received by any of the other candidates. Thus, expected plurality maximizing candidates are assumed to not care about beating alternatives which can not win the election anyway, such as abstention. This implicitly rules out nonstrategic alternatives which can win the election, such as a choice of "None of the above," for instance. An expected plurality maximizing candidate is also not completely insensitive to the distribution of the votes she does not receive.

2.3.3 Maximizing Probability of Victory

Given opponents' pure strategies x_{-j}, a *probability of victory maximizing candidate* $j \in \mathcal{J}^0$ seeks to maximize

$$R_j(x) = \sum_{a \in A} \left(\frac{1}{|W(a)|} \mathbf{1}[j \in W(a)] \Pr[a] \right). \tag{2.3}$$

2.3.4 Payoffs with Mixed Strategies

We now extend the above objective functions to the case where candidates may use mixed strategies. This is a standard exercise, and is necessary only because our extension implies that candidates possess von Neumann-Morgenstern payoff functions regardless of their objective.

Given any candidate objective function $u_j : Y \to \mathbb{R}$, we will write

$$u_j(x_j; \mu_{-j}) = \int_X \cdots \int_X u_j(x) \mu_1(dx_1) \ldots \mu_{j-1}(dx_{j-1}) \mu_{j+1}(dx_{j+1}) \ldots \mu_{J_0}(dx_{J_0})$$

for the expected value to candidate j of action x_j by candidate j, given mixed strategies $\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_{J_0}$ by the other candidates.

Similarly, given candidate objective function $u_j : Y \to \mathbb{R}$, we will write

$$u_j(\mu_j; \mu_{-j}) = u_j(\mu) = \int_X \cdots \int_X u_j(x) \mu_1(dx_1) \ldots \mu_{J_0}(dx_{J_0})$$

for the expected value to candidate j of mixed strategy μ_j by candidate j, given mixed strategies $\mu_1, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_{J_0}$ by the other candidates.

2.4 Two Different Definitions of Equivalence

There are at least two definitions of equivalent objective function, best response equivalence and equilibrium equivalence. The first, and most demanding, is *best response equivalence.* In words, best response equivalence holds whenever two objective functions prescribe identical optimal strategies regardless of the strategy chosen by the *opponent(s)*. Best response equivalence may hold with respect to all strategies or only with respect to pure strategies.[1] Such equivalence is essentially a decision-theoretic concern, as the strategic effects of other players' motivations are inconsequential to the player in question.

Formally, we define best response equivalence in pure strategies as

Definition 2 *Two payoff functions u_i and u'_i exhibit* best response equivalence in pure strategies *for candidate i if, for all $x_{-i} \in Y_{-i}$,*

$$\arg\max_{x \in X_i} u_i(x, x_{-i}) = \arg\max_{x \in X_i} u'_i(x, x_{-i}).$$

Similarly, we define best response equivalence in mixed strategies as

Definition 3 *Two payoff functions u_i and u'_i exhibit* best response equivalence in mixed strategies *for candidate i if, for all $\mu_{-i} \in M_{-i}$,*

$$\arg\max_{\mu_i \in M_i} u_i(\mu_i, \mu_{-i}) = \arg\max_{\mu_i \in M_i} u'_i(\mu_i, \mu_{-i}).$$

A second, and weaker, form of equivalence, equilibrium equivalence, holds whenever the set of Nash equilibria under two different objective functions are identical. Just as with best response equivalence, we can speak of equilibrium equivalence holding with respect to the space of all strategies, or just with respect to the set of pure strategy equilibria.[2]

Formally, we define the two types of equilibrium equivalence as follows.

Definition 4 *Two vectors of payoff functions u and u' exhibit* equilibrium equivalence in pure strategies *if for all $x \in Y$,*

$$x \in PNE(u) \Leftrightarrow x \in PNE(u').$$

[1] Since players possess von Neumann-Morgenstern utility functions, then mixed strategy best response equivalence implies pure strategy best response equivalence, but the reverse implication does not hold, as we show in Section 2.5.2.

[2] As with best response equivalence, equilibrium equivalence in mixed strategies implies equilibrium equivalence in pure strategies, but the reverse implication does not necessarily hold.

Definition 5 *Two vectors of payoff functions u and u' exhibit* equilibrium equivalence in mixed strategies *if for all $\mu \in M$,*

$$\mu \in NE(u) \Leftrightarrow \mu \in NE(u').$$

Obviously, both of these definitions are satisfied whenever best response equivalence in mixed strategies holds.[3] For completeness, we prove this formally.

Theorem 6 *Let $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$ and $\Gamma' = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u')$. If u and u' exhibit best response equivalence in mixed strategies, then u and u' exhibit equilibrium equivalence in mixed strategies.*

Proof: Suppose, by way of contradiction, that u and u' exhibit best response equivalence in mixed strategies but fail to exhibit equilibrium equivalence in mixed strategies. Then there must exist a profile, μ^*, which is an equilibrium under one set of objective functions, say u, but which is not an equilibrium under the other set, u'. Then there exists some candidate k for whom, given μ^*_{-k} and maximization of u'_k, there exists a unilateral deviation, $\hat{\mu}_k \neq \mu^*_k$, such that

$$u'_k(\hat{\mu}_k; \mu^*_{-k}) > u'_k(\mu^*).$$

Then it is the case that the best response correspondence for candidate k under objective function u' does not contain μ^*_k when $\mu_{-k} = \mu^*_{-k}$. However, since μ^* is an equilibrium under u, then the best response correspondence for candidate k under objective function u must contain μ^*_k when $\mu_{-k} = \mu^*_{-k}$.

Thus, for candidate K facing opponents' mixed strategy profile μ^*_{-k},

$$\arg \max_{\mu_k \in M_k} u_k(\mu_k, \mu^*_{-k}) \neq \arg \max_{\mu_k \in M_k} u'_k(\mu_k, \mu^*_{-k}),$$

contradicting the fact that u and u' exhibit best response equivalence in mixed strategies.

[3]For a more detailed discussion of this, see Aranson, Hinich, and Ordeshook (1974), p. 144-145.

Since supposing that equilibrium equivalence in mixed strategies does not hold when best response equivalence in mixed strategies does hold leads to a contradiction, it must be the case that best response equivalence in mixed strategies implies equilibrium equivalence in mixed strategies. ∎

The next corollary follows immediately from Theorem 6. It states that best response equivalence in pure strategies implies equilibrium equivalence in pure strategies.

Corollary 7 *Let* $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$ *and* $\Gamma' = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u')$. *If* u *and* u' *exhibit best response equivalence in pure strategies, then* u *and* u' *exhibit equilibrium equivalence in pure strategies.*

Of course, the converses of Theorem 6 and Corollary 6 do not hold in general. Nevertheless, while equilibrium equivalence is a weaker criterion for equivalence, it is often "enough" for our purposes, since most analyses of electoral competition are solely concerned with (possibly some refinement of) Nash equilibrium.

We now examine the question of when best response equivalence does (and does not) hold between expected plurality, expected vote share, and probability of victory.

2.5 Best Response Equivalence

Spatial models of elections often assume that the candidates' sole goal is victory. To calculate the optimal strategy for such a candidate, one must take into account the probability of victory resulting from each strategy. In general, this probability is not a trivial computation, especially when studying probabilistic voting models (e.g., Hinich (1977), Coughlin and Nitzan (1981a), (1981b), Ledyard (1984), and Chapter 3 of this work). For this reason, researchers have sought candidate objectives which are easier to compute and yet lead to the same predictions as those generated by probability of victory.

The existing literature has shown that equivalence results do not hold in general. In order to show equivalence between maximizing plurality and probability of victory, three things are assumed to hold. Voters choices are mutually independent conditional upon the policy choices of the candidates, all voters, conditional upon any policy profile x, have identical expected behavior, and candidate's strategies are announced simultaneously. We label the first two requirements *independence* and *symmetry*, respectively, and we describe any electoral game which satisfies these requirements as *possessing a representative voter*, since the expected behavior of any voter can be inferred from the expected behavior of any other voter.

In Section 2.5.1 we prove our main result, Theorem 13, which states that maximization of expected plurality, maximization of expected vote, and maximization of probability of victory exhibit best response equivalence in any electoral game satisfying the following conditions: the electorate is finite, there are two candidates, abstention is not allowed, and the game possesses a representative voter. In Section 2.5.3 we examine the issue of abstention in two candidate elections, the case examined by Ledyard. An example is provided which shows that Theorem 13 can not be extended to this case. Section 2.5.4 contains Hinich's claim and a counterexample. Section 2.5.5 contains a discussion of elections involving more than 2 candidates, including an example showing that asymptotic best response equivalence between expected vote share and probability of victory is not generally true in such elections.

2.5.1 2 Candidates, No Abstention

First it is shown that, for all N with $J_0 = J = 2$, maximizing expected plurality, π_j, is equivalent to maximizing expected vote share, V_j.

Proposition 8 *Assume that $J = 2$ and Assumption 1 holds. Then, for any j and all x_{-j},*

$$\arg\max_{x_j \in X} \pi_j(x_j; x_{-j}) = \arg\max_{x_j \in X} V_j(x_j; x_{-j})$$

Proof: By Assumption 1 and $J = 2$,

$$\begin{aligned}\pi_1(x) &= \frac{1}{N}\sum_{i=1}^{N}[p_i^1(x) - p_i^2(x)] \\ &= \frac{1}{N}\sum_{i=1}^{N}[2p_i^1(x) - 1] \\ &= \frac{2}{N}[\sum_{i=1}^{N} p_i^1(x)] - 1 \\ &= 2V_j(x) - 1.\end{aligned}$$

Thus, since the choice of candidate 1 is arbitrary, π_j is an increasing affine transformation of V_j, proving the proposition. ∎

Representative Voter

We now restrict attention to elections which satisfy an admittedly stringent symmetry condition. In particular, we require for all voters to have identical response functions. Formally, we make the following assumption.

Assumption 9 (Symmetry) *For all $i, j \in \mathcal{N}$ and all $x \in Y$,*

$$p_i(x) = p_j(x).$$

As mentioned above, we describe any game with $J = J^0 = 2$ which satisfies Assumptions 1 and 9 as *possessing a representative voter*, since the entire electorate's expected behavior can be expressed as a function of a single voter's expected behavior. In particular, we can drop the subscripts from the voters' response functions, as they are identical to one another. We formally define this class of games below.

Definition 10 *Any electoral game, $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$, in which the set of voters, $\mathcal{N}$, is finite and the set of response functions, p, satisfies Assumptions 1 and 9 is said to* possess a representative voter.

We now show that, in any electoral game with a finite electorate and possessing a representative voter, maximization of expected vote share, maximization of expected plurality, and maximization of probability of victory exhibit best response equivalence in pure strategies.

The following lemmata make the proof of Theorem 13 almost immediate. In particular, Lemma 11 allows us to use differential calculus to show that the probability of victory is a strictly increasing function of expected vote when an electoral game with 2 candidates and no abstention possesses a representative voter. In a more general sense, Lemma 11 ensures that the probability of victory is a smooth function of the vector of all voters' response functions.

Lemma 11 *Let $\mathcal{N}$ and $\mathcal{J}$ be finite sets. If Assumption 1 holds, then $R_j(x)$ is continuously differentiable in $p(x)$ for all j, for arbitrary values of J.*

Proof: Fix a vector of policy proposals $x \in Y$. The probability of a vector of votes $\hat{a} \in A$ is given by

$$f(a; p(x)) = \prod_{i=1}^{N} p_i^{a_i}(x),$$

where a_i denotes the alternative receiving i's vote whenever a is the vector of votes. The probability of alternative j receiving the most votes is

$$R_j(x) = \sum_{a \in A} \frac{1}{|W(a)|} \left[\mathbf{1}[j \in W(a)] f(a; p(x)) \right].$$

For any $a \in A$, $x \in Y$, and $i \in \mathcal{N}$, $f(a; p(x))$ is linear, and hence continuously differentiable, in p_i. Since R_j is a linear combination of $f(a; p(x))$ for different values of a, it follows that R_j is continuously differentiable in $p(x)$. ∎

The next result, Lemma 12, states that the conditional expected value of the upper 50% tail of the binomial distribution is strictly positive. We use this fact to show that the probability of victory is a strictly increasing function of expected vote share.

Lemma 12 *For any positive integer N,*

$$\sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] \geq 0, \tag{2.4}$$

where the inequality is strict for all $p \in (0,1)$.

Proof: Let X be a random variable distributed according to a Binomial(N, p) distribution. Let $Z = X - Np$ denote the mean zero standardization of X. It is obvious that if $p = 0$ or $p = 1$ then Equation 2.4 is satisfied with equality. Therefore, assume $p \in (0,1)$.

First, supposing that $p \in (0,1)$, the expected value of Z can be rewritten as

$$\begin{aligned} \sum_{c=0}^{N} \binom{N}{c} \left[p^{c}(1-p)^{N-c}(c-Np)\right] &= 0, \\ p(1-p) \cdot \sum_{c=0}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] &= 0, \end{aligned}$$

so that, dividing through by $p(1-p)$, we obtain

$$\sum_{c=0}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] = 0. \tag{2.5}$$

Next, notice that, supposing that $p \in (0,1)$, for all $c < Np$,

$$\binom{N}{c} p^{c}(1-p)^{N-c}(c-Np) < 0,$$

while, for all $c > Np$,

$$\binom{N}{c} p^{c}(1-p)^{N-c}(c-Np) > 0,$$

and, for $c = Np$,

$$\binom{N}{c} p^{c}(1-p)^{N-c}(c-Np) = 0.$$

Accordingly, supposing that $p \in (0,1)$, for all $c < Np$,

$$\begin{aligned} \binom{N}{c} p^c (1-p)^{N-c}(c-Np) &< 0, \\ p(1-p)\cdot\binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &< 0, \\ \binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &< 0, \end{aligned} \tag{2.6}$$

while, for all $c > Np$,

$$\begin{aligned} \binom{N}{c} p^c (1-p)^{N-c}(c-Np) &> 0, \\ p(1-p)\cdot\binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &> 0, \\ \binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &> 0, \end{aligned}$$

and, for $c = Np$,

$$\begin{aligned} \binom{N}{c} p^c (1-p)^{N-c}(c-Np) &= 0, \\ p(1-p)\cdot\binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &= 0, \\ \binom{N}{c} p^{c-1}(1-p)^{N-c-1}(c-Np) &= 0. \end{aligned}$$

We complete the proof in two cases. The first case is when $0 < p \leq \frac{1}{2}$. Combining Equations 2.5 and 2.6, we see that

$$\sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right]$$

is a sum of nonnegative (and at least one strictly positive) terms whenever $p \leq \frac{1}{2}$, implying the result in this case.

In the final case, $\frac{1}{2} < p < 1$, the sum

$$\sum_{0}^{c=\lfloor \frac{N}{2} \rfloor} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right]$$

is completely comprised of strictly negative terms. Rewriting this sum as

$$\sum_{c=0}^{N} \binom{N}{c} \left[p^{c}(1-p)^{N-c}(c-Np)\right] \quad - \quad \sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] < 0$$

$$0 \quad - \quad \sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] < 0,$$

implying that

$$\sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p^{c-1}(1-p)^{N-c-1}(c-Np)\right] > 0$$

as desired. ∎

Using Lemmas 11 and 12, we can now prove the main result in this section. In any 2 candidate election without abstention, a sufficient condition for expected plurality, expected vote share, and probability of victory to be equivalent is that the electoral game possess a representative voter.

Theorem 13 *For any game possessing a representative voter,*

$$x_j \in \arg\max_X V_j(x_j; x_{-j}) \Leftrightarrow x_j \in \arg\max_X R_j(x_j; x_{-j}).$$

Proof: Since $V_j(x_j; x_{-j})$ is a strictly increasing function of $p^j(x)$, it suffices to show that $R_j(x_j; x_{-j})$ is also an increasing function of $p^j(x)$.

Lemma 11 ensures that we can differentiate $R_j(x_j; x_{-j})$ with respect to $p^j(x)$. For notational ease, let $R_j = R_j(x_j; x_{-j})$ and $p_j = p^j(x)$. Taking the first derivative of

R_j with respect to p_j, we obtain

$$\begin{aligned}
\frac{\partial R_j}{\partial p_j} &= \sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[c p_j^{c-1}(1-p_j)^{N-c} - (N-c)p_j^c(1-p_j)^{N-c-1} \right], \\
&= \sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p_j^{c-1}(1-p_j)^{N-c-1} \left(c(1-p_j) - (N-c)p_j \right) \right], \\
&= \sum_{c=\lceil \frac{N}{2} \rceil}^{N} \binom{N}{c} \left[p_j^{c-1}(1-p_j)^{N-c-1}(c - Np_j) \right], \qquad (2.7) \\
&> 0, \qquad (2.8)
\end{aligned}$$

where the final inequality comes from Lemma 12. Thus, the probability of victory is a strictly increasing function of the expected vote and by Proposition 8, a strictly increasing function of expected plurality. ∎

Theorem 13 only gives sufficient conditions for best response equivalence. The following examples show that Assumptions 1 and 9, respectively, are not necessary for best response equivalence.

Example 14 In this example, the individual response functions violate Assumption 1. In particular, realizations of a_i and a_j are perfectly correlated for all $i, j \in \mathcal{N}$. We show that, nevertheless, best response equivalence holds between all three objective functions.

Let $X = \{L, R\}$ be a binary policy space, $\mathcal{J} = \mathcal{J}^0 = \{1, 2\}$, and $N = 3$. The voters' response functions are identical, but do not satisfy Assumption 1. In particular, the voters' responses are given by the following rule, where a_i denotes the action of voter i, and a denotes the vector of all a_i.

$$a = \begin{cases} (1,1,1) & \text{if} \quad x = (L, R) \\ \begin{cases} (1,1,1) & \text{with probability } 0.5 \\ (2,2,2) & \text{with probability } 0.5 \end{cases} & \text{if} \quad x = (L, L) \\ \begin{cases} (1,1,1) & \text{with probability } 0.5 \\ (2,2,2) & \text{with probability } 0.5 \end{cases} & \text{if} \quad x = (R, R) \\ (2,2,2) & \text{if} \quad x = (R, L) \end{cases}$$

That is, in all states, the voters vote unanimously for one candidate, and prefer position L.

Regardless of whether a candidate is maximizing expected vote share, expected plurality, or probability of victory, the pure strategy L weakly dominates all other pure and mixed strategies. In fact, L is a best response for either candidate to any strategy chosen by her opponent under any of the three objective functions. Thus, best response equivalence holds in this case, even though Assumption 1 does not hold. $\triangle$

Example 15 In this example, the response functions do not satisfy Assumption 9. In particular, two voters' behaviors are the "mirror image" of each other, while the third voter's behavior is invariant to the policy chosen by the candidate. We show that best response equivalence still holds between all three objective functions in this example, even though the game does not possess a representative voter.

Let $X = [0,1]$, $N = 3$, $\mathcal{J} = \mathcal{J}^0 = \{1, 2\}$, and let the voters' response functions satisfy Assumption 1. In particular, assume the following response functions.

$$\begin{aligned} p_1^1(x) &= \frac{1}{2} \\ p_2^1(x) &= \frac{1}{2} + \frac{1}{2}(x_1 - x_2) \\ p_3^1(x) &= \frac{1}{2} - \frac{1}{2}(x_1 - x_2). \end{aligned}$$

Thus, voter 1's behavior is completely unresponsive to the policies announced by the candidates, while voters 2 and 3 each are more likely to choose the candidates announcing the rightmost and leftmost policies, respectively.

It follows easily that the expected vote share and expected plurality of either candidate is invariant to the vector of policies chosen by the candidates, with each candidate receiving an expected vote share of 0.5 and an expected plurality of zero. In calculating the best response correspondence for candidate 2 under maximization of probability of victory, we obtain

$$\begin{aligned} R_2(x) &= (0.5)\left[1 - 2[x_1 - x_2]^2 + 2[x_1 - x_2]^2\right] \\ &= 0.5 \end{aligned}$$

for all choices of x_1 and x_2, implying that a probability of victory maximizing candidate is indifferent between all policies, regardless of the opponent's strategy. Since this holds under all of the three objective functions, best response equivalence holds in this model, in which voters' behavior does not satisfy Assumption 9. $\triangle$

2.5.2 Best Response Equivalence in Mixed Strategies

We now show that even if Assumptions 1 and 9 are satisfied in a 2 candidate election without abstention, best response equivalence in mixed strategies does not necessarily hold. In particular, it is possible for a candidate to be indifferent between all possible pure actions under one objective function and not under another, implying that the best response correspondence under the first objective function contains at least one totally mixed strategy while the best response correspondence under the second objective function contains no such mixed strategies.

Example 16 This example shows that Let $X = \{\alpha, \beta, \gamma\}$, $\mathcal{J} = \mathcal{J}_0 = \{1, 2\}$, $\mathcal{N}$ be finite, and assume that the game possesses a representative voter.

Suppose that $p(x_1, x_2)$ is given by

	α	β	γ
α	(0.3,0.7)	(0.3,0.7)	(0.3,0.7)
β	(0.3,0.7)	(0.3,0.7)	(0.3,0.7)
γ	(0.9,0.1)	(0,1)	(0,1)

Table 2.1: Expected Vote

First consider the case where $N = 1$. In this case, the probability of victory for candidate one and expected vote share are identical, so that best response equivalence holds trivially. Now consider the case where N is arbitrarily large. In this case, given Assumptions 1 and 9, the payoffs of an expected vote share maximizing candidate are still given by Table 16. However, probability of victory maximizing candidates face payoffs given, for arbitrarily small $\varepsilon > \delta > 0$,

	α	β	γ
α	$(\varepsilon, 1-\varepsilon)$	$(\varepsilon, 1-\varepsilon)$	$(\varepsilon, 1-\varepsilon)$
β	$(\varepsilon, 1-\varepsilon)$	$(\varepsilon, 1-\varepsilon)$	$(\varepsilon, 1-\varepsilon)$
γ	$(1-\delta, 1)$	0	0

Table 2.2: Probability Of Victory, N Large

Now suppose that candidate 2 is mixing with equal probability between the three policies, α, β, and γ. Then the expected vote share offered to candidate 1 by any of the three positions is equal to 0.3. However, the probability of victory offered by α and β is approaching 0, while the probability of victory offered by γ is approaching $\frac{1}{3}$. Thus, in large elections, the best response correspondence of an expected vote share maximizing candidate would contain all mixtures over X[4], while that of a probability of victory maximizing candidate would only contain the pure strategy $x_1 = \gamma$. △

Aside from its obvious role as a counterexample to an extension of Theorem 13, Example 16 is also interesting for the following reason. Hinich's and Ledyard's results regarding asymptotic best response equivalence in pure strategies rely on appeals to

[4]It can be easily verified that the same is true for an expected plurality maximizing candidate: simply subtract $\frac{1}{2}$ from every payoff for both candidates in Table 16.

forms of the Law of Large Numbers, while Example 16 shows that the same logic actually breaks equivalence in mixed strategies.

2.5.3 2 Candidates with Abstention

The previous section provided a theorem which strengthens Hinich's statement that, in two candidate elections without abstention and without coordination by voters, maximizing plurality and probability of victory yielded equivalent strategies in equilibrium. As discussed in the conclusion, the proposition proved here for $J = 2$ is both weaker and stronger than Hinich's original claim, but it is obviously concerned with a very special case, since abstention is generally allowed in most elections, for example. When abstention is allowed, maximizing expected vote is generally not equivalent to maximizing plurality, as we show in Example 19. In this section, we provide an example of a 2 candidate election in which abstention is allowed and voters' behavior satisfies both Assumptions 1 and 9, but maximizing plurality and maximizing probability of victory do not exhibit best response equivalence.

Ledyard's Result

Ledyard (1984) provides a result stating that, when the number of voters is large enough, maximization of $V(x)$ and $R_j(x)$ are equivalent when $J^0 = 2$ and $J = 3$. For clarity, we quote the claim, replacing Ledyard's notation with the concepts they denote in his model.

> If [the number of voters] is large, then [expected plurality] is a good approximation for a candidate to use in place of [probability of victory].
>
> ... Since the [individual vote choices] are independently identically distributed, it follows from a Law of Large Numbers that ... maximizing [Candidate A's expected plurality] maximizes (in the limit) the probability that A wins. [Ledyard, (1984), pp. 20-21.]

Ledyard proves his result at the limit (i.e., an infinite number of voters), which is never realized in his model. As the quote makes clear, Ledyard's argument is *that expected* plurality is a good approximation of probability of victory in large electorates. We are silent on this issue. What we now show is that the abstention can not be allowed in order for Theorem 13 to hold. The next example, due to John Duggan, highlights why best response equivalence may fail to hold in 2 candidate elections with abstention, *even when the election possess a representative voter.*

Example 17 This example shows how, even in a two candidate election, a significant change in expected total turnout can alter the candidates' probabilities of victory without altering their expected pluralities.

We assume that voter behavior satisfies Assumptions 1 and 9. Let $N = 3$, $J = 3$, $J^0 = 2$, and consider two policy positions, $x, y \in Y$, with $x = (x_1, x_2)$ and $y = (x'_1, x_2)$, characterized by the following voter behavior, where $p(z) = (a, b, c)$ means that, given policy proposal vector z, the probability of any given voter voting for candidate 1 is a, while the probability of voting for candidate 2 is b and the probability of abstention is c:

$$\begin{aligned} p(x) &= (0.08, 0.02, 0.90) \text{ and} \\ p(y) &= (0.53, 0.47, 0.00). \end{aligned}$$

We focus on candidate 1. It is straight forward to compute the following:

$$\begin{aligned} \pi_1(x) &= 0.0599\bar{3}\bar{3} \\ R_1(x) &= 0.581396, \end{aligned}$$

while

$$\begin{aligned} \pi_1(y) &= 0.06 \\ R_1(y) &= 0.544946, \end{aligned}$$

so that $\pi_1(x) < \pi_1(y)$, but $R_1(x) > R_1(y)$.

The reason that the two objective functions are not equivalent is that, conditional on any given voter showing up, the probability of candidate 1 receiving that voter's vote is much higher at x than it is at y. It is interesting to note that if x_1 and x'_1 are her only choices, candidate 1 has a strict incentive to reduce expected turnout if she wishes to maximize her probability of victory. △

2.5.4 Hinich's Result

For over twenty years, the theoretical literature has been largely silent on the implications of the modeler's choice of candidates' objectives. We suggest that a re-examination of this silence is necessary.

First, the validity of the claim in Hinich (1977) regarding asymptotic equivalence of maximizing expected vote and maximizing probability of victory in two candidate elections is not obvious. For clarity, we quote the claim.

> If voters in a large electorate act independently, the distribution of a candidate's total vote approximates a normal distribution for Bernoulli trials. The mean of this normal distribution is the expected vote. Thus for large electorates, maximizing probability of victory is equivalent to maximizing expected vote, *which is also equivalent to maximizing plurality since everyone votes.* [Hinich (1977), pp. 212-213, Italics in original.]

This claim has been cited by several authors (including Coughlin and Nitzan (1981a), (1981b), Enelow and Hinich (1989), and Ledyard (1984), among others). Exactly when Hinich's claim holds is an open question, however. Theorem 13 states that Hinich's claim is correct for finite electorates whenever the game possesses a representative voter. We now show, however, that it is not the case that best response equivalence holds in all 2 candidate elections without abstention. In particular, we construct an example in which voters' behavior does not satisfy Assumption 9 and best response equivalence does not hold in *any* finite electorate, contradicting Hinich's result.

Example 18 This example highlights the fact that, even in a 2 candidate election without abstention, best response equivalence between expected plurality and probability of victory can fail if voters do not satisfy our symmetry assumption, Assumption 9.

Let $J = J^0 = 2$ and $N = 3$. Consider an election in which $X = \{L, R\}$ and voter behavior is given by

$$p_1^1(L,L) = p_2^1(L,L) = p_3^1(L,L) = \tfrac{1}{2}$$
$$p_1^1(R,R) = p_2^1(R,R) = p_3^1(R,R) = \tfrac{1}{2},$$

while

$$\begin{aligned} p_1^1(R,L) &= \varepsilon \\ p_2^1(R,L) &= \frac{3}{4} - \varepsilon \\ p_3^1(R,L) &= \frac{3}{4} - \varepsilon, \end{aligned}$$

and

$$\begin{aligned} p_1^1(L,R) &= 1 - \varepsilon \\ p_2^1(L,R) &= \frac{1}{4} + \varepsilon \\ p_3^1(L,R) &= \frac{1}{4} + \varepsilon \end{aligned}$$

Consider the pure strategy profile (L, L). Note that this is a strict Nash equilibrium under maximization of expected vote share:

$$V_1(L,L) = V_2(L,L) = \frac{1}{2},$$

while

$$\begin{aligned}V_1(R,L) &= \frac{1}{3}(\varepsilon+\frac{3}{4}-\varepsilon+\frac{3}{4}-\varepsilon)\\ &= \frac{1}{3}(\frac{3}{2}-\varepsilon)\\ &= \frac{1}{2}+\frac{\varepsilon}{3}\\ &< \frac{1}{2}\\ V_1(R,L) &< V_1(L,L),\end{aligned}$$

and, by the symmetry of the voters' behavior,

$$\begin{aligned}V_2(L,R) &= \frac{1}{3}(\varepsilon+\frac{3}{4}-\varepsilon+\frac{3}{4}-\varepsilon)\\ &= \frac{1}{2}-\frac{\varepsilon}{3}\\ &< \frac{1}{2}\\ V_2(L,R) &< V_2(L,L).\end{aligned}$$

Thus, by Proposition 8, (L, L) is also an equilibrium under maximization of expected plurality. Note that (L, L) is a Nash equilibrium under these objective functions for any $\varepsilon \in [0, \frac{1}{4}]$.

Now suppose that candidate 1 deviates to R, both her expected vote and expected plurality decrease by ε, but for her probability of victory is

$$\begin{aligned}R_1(R,L) &= \varepsilon(\frac{3}{4}-\varepsilon)(\frac{3}{4}-\varepsilon)+(1-\varepsilon)(\frac{3}{4}-\varepsilon)(\frac{3}{4}-\varepsilon)\\ &\quad +\varepsilon(\frac{3}{4}-\varepsilon)(\frac{1}{4}-\varepsilon)+\varepsilon(\frac{1}{4}-\varepsilon)(\frac{3}{4}-\varepsilon)\\ &= (\frac{3}{4}-\varepsilon)^2+2\varepsilon(\frac{3}{4}-\varepsilon)(\frac{1}{4}-\varepsilon)\\ &= \frac{9}{16}-\frac{3}{2}\varepsilon+\varepsilon^2+2\varepsilon(\frac{3}{16}-\varepsilon+\varepsilon^2)\\ &= \frac{9}{16}-\frac{30}{16}\varepsilon-\varepsilon^2+2\varepsilon^3.\end{aligned}$$

Given $\varepsilon > 0$, the difference between candidate 1's probability of victory at (R, L) and her probability of victory at (L, L) is

$$\begin{aligned} R_1(R,L) - R_1(L,L) &= \frac{9}{16} - \frac{30}{16}\varepsilon - \varepsilon^2 + 2\varepsilon^3 - \frac{1}{2} \\ &= \frac{1}{16} - \frac{30}{16}\varepsilon - \varepsilon^2 + 2\varepsilon^3. \end{aligned}$$

Now, letting ε get arbitrarily close to zero, we see that

$$\begin{aligned} \lim_{\varepsilon\to 0} [R_1(R,L) - R_1(L,L)] &= \lim_{\varepsilon\to 0} \left[\frac{1}{16} - \frac{30}{16}\varepsilon - \varepsilon^2 + 2\varepsilon^3\right] \\ &= \frac{1}{16} \\ &> 0, \end{aligned}$$

so that, for sufficiently small positive values of ε, L is not a best response for candidate 1 if candidate 2 chooses L. In other words, for sufficiently small positive values of ε, (L, L) is a Nash equilibrium under maximization of expected vote share or expected plurality, but *not* a Nash equilibrium under maximization of probability of victory.

This example can be extended to arbitrary numbers of voters. Assuming that $J = 2$ and N odd, assume that $p_i^1(L, L) = \frac{1}{2}$ for all i, and take voter behavior to be such that

$$\begin{aligned} p_i^1(R,L) &= \varepsilon & \forall i \le \tfrac{N-1}{2}, \\ p_i^1(R,L) &= \tfrac{N}{N+1} - \varepsilon & \forall i > \tfrac{N-1}{2}. \end{aligned}$$

Again, deviating from (L, L) to (R, L) decreases both candidate 1's expected vote share and expected plurality by ε, but increases her probability of victory. Indeed, denoting candidate 1's probability of victory by $R_1(\varepsilon, N)$, it can be shown that

$$\lim_{\varepsilon\to 0}\lim_{N\to\infty} R_1(\varepsilon, N) = \frac{1}{\sqrt{e}} > \frac{1}{2}.$$

Thus, *even asymptotically*, best response equivalence between the two objective functions does not hold in general when voters' behavior fails to satisfy Assumption 9.

It should be noted that this example can be trivially extended to a larger policy *space (including* an interval of $\mathbb{R}$. To see this, just let L and R denote the endpoints of the interval, and give voters' linear response functions. It is easily verified that the remainder of the analysis remains the same: (L, L) remains a Nash equilibrium vote or plurality maximizing candidate while R is the best response to L for a probability of victory maximizing candidate 1. $\triangle$

Note, as a note for our discussions later in this Chapter, that Example 18 also provides an example of a case where, in a 2 candidate election without abstention, probability of victory and expected plurality maximization do not even exhibit equilibrium equivalence.

2.5.5 More Than 2 Candidates

Another open question regards elections between more than two candidates. What can we say, if anything, about the relationship between different incentives in such a framework? The next example, due to Tom Palfrey, shows that asymptotic equivalence in our framework is not possible without more restrictions.

Example 19 This example utilizes the fact that an increase in one candidate's expected vote share does not necessarily imply a decrease in every other candidate's expected vote share.

Let the policy space be the unit interval, voters' preferences be Euclidean, and let there be three candidates. We assume that voter behavior is sincere: voters vote for the candidate whose announced position is closest to the voter's ideal policy. We also assume there is a continuum of voters. Assume that candidates 1 and 2, proposing x and y, respectively, are adopting identical strategies. Candidate 3 is adopting a strategy, z, which is different from that chosen by candidates 1 and 2. As it stands now, candidates 1 and 2 are each receiving 45% of the vote, while candidate 3 is receiving only 10% of the vote. The probability of victory for candidates 1 and 2 is also equal at this strategy profile. In particular, each candidate wins half of the time, while candidate 3 never wins.

Assume now that candidate 1 seeks to maximize her expected vote share and that there exists a policy x' such that, given y and z, she will receive 47% of the vote at x', candidate 3 will receive 5% of the vote, and candidate 2 will receive 48% of the vote. Thus, candidate 1's vote share has increased, but her probability of victory has gone to zero, since candidate 2's vote share is higher than candidate 1's.

This example shows that, even with a continuum of voters, a candidate may increase her expected vote share but decrease her probability of victory. This is generally the case when there are more than 2 candidates. △

Notice that Example 19 does not violate equivalence between expected plurality, as defined above, and probability of victory. The question of asymptotic equivalence between expected plurality and probability of victory with more than two candidates is an open question, one which we discuss to some degree later in this chapter.

2.6 Equilibrium Equivalence

2.6.1 Introduction

We now extend the study of candidate objective functions to the question of equilibrium equivalence. As discussed earlier, best response equivalence is essentially a decision-theoretic concern, as it is defined to hold regardless of the opponents' strategies. Equilibrium equivalence, on the other hand, is a game-theoretic concern. Two objective functions are said to exhibit equilibrium equivalence if the sets of Nash equilibria under the two objective functions are identical.

In this section, we examine strict local Nash equilibria - profiles of candidate strategies in which no candidate has an arbitrarily small deviation which leads to a weakly greater payoff. Our motivation for examining local equilibria is two-fold. Obviously, every global equilibrium is also a local equilibrium. If local equilibrium equivalence fails at all local equilibria, then global equilibrium equivalence fails, implying by Theorem 6 that best response equivalence must fail as well. The following section provides a result which states that, for generic continuously differentiable *voter response* functions, the first order conditions for local equilibrium equivalence fail to be satisfied, implying that, generically, best response equivalence fails to hold between maximization of expected plurality and probability of victory maximization.

Secondly, the global equilibrium constructed in Theorem 46 in Chapter 3 is, intuitively, a local equilibrium which "becomes" global. Hence, it is hoped that an understanding of the general properties of local equilibrium equivalence will simultaneously lead to an understanding of the robustness of the equilibrium constructed in Chapter 3 to the specification of the candidates' objective functions.

Our first result in this section is that interior p-symmetric strict local equilibria under maximization of expected vote share and maximization of probability of victory are identical whenever voters' types are independently distributed and the second derivative of each candidates' expected vote share is negative definite at the local equilibrium. We show that local equilibria under maximization of probability of victory are also local equilibria under maximization of expected vote share even if we relax the local negative definiteness of the second derivative of each candidate's expected vote share to local strict concavity of each candidate's expected vote share with respect to her own policy choice.

These positive local equilibrium equivalence results are motivated by the results of several previous papers in probabilistic voting models of candidate competition. For instance, Coughlin and Nitzan (1981a), (1981b) examine local Nash equilibria for two candidate elections under a probabilistic voting model, and our results apply in much of their framework, as they require voters "make independent voting decisions," which corresponds to Assumption 1 in our framework. Additionally, our equilibrium equivalence results apply to the framework studied by Hinich (1977).

Similarly, in Chapter 3, we examine a model of strategic probabilistic voting with an arbitrary number of candidates seeking to maximize expected margin of victory. We prove the existence of a p-symmetric strict Nash equilibrium at the point that maximizes the sum of the voters' utility functions whenever the number of voters is large enough. Our method of proof, using a Taylor Series expansion of the expected payoffs of a deviation by any candidate), utilizes the fact that the point which max-

imizes the sum of voters' utilities is a local critical point which becomes a global expected payoff maximum as the number of voters grows without bound.

Finally, we prove that interior p-asymmetric local equilibria are generically not equivalent under maximization of expected vote share and maximization of probability of victory whenever voters' behaviors are continuously differentiable. We show this by noting that the first order conditions for interior equilibrium are almost always not met simultaneously under the two objective functions.

2.6.2 The Model

Let $\mathcal{J}$, with $|\mathcal{J}| = J$, denote the set of candidates, as we do not deal with abstention in this section, and let $\mathcal{N}$, with $|\mathcal{N}| = N$, denote the set of voters. Each candidate simultaneously chooses a point in some policy space X. We denote the space of all J-dimensional vectors of policy proposals by Y.

We write the action of voter i, given $y \in Y$, as $a_i \in \mathcal{J}$, and denote the number of votes received by candidate j by $v_j = |\{i \in \mathcal{N} | a_i = j\}|$. We write a for the vector of a_i for all voters i and A for the space of all such vectors. Again, as above, we denote the probability that voter i votes for candidate j at $y \in Y$ by $p_i^j(y)$ and the vector of all $p_i^j(y)$, for some candidate j and all voters i, by $p^j(y)$. We will note any additional assumption we make about p_i as we need them. We continue to refer to p_i as a *response function* for voter i.

For any $x \in Y$, let $G_j(i, x)$ denote the probability candidate j wins, conditional on voter i voting for j (i.e., $a_i = j$) and let $H_j(i, x)$ denote the conditional probability that candidate j wins, conditional on $a_i \neq j$. Formally,

$$\begin{aligned} G_j(i,x) &= \Pr\{w(a) = j | p_{-i}(x); a_i = j\} \\ &= \sum_{a \in A} \left[\frac{1}{|W(a)|} \Pr\{a | p_{-i}(x); a_i = j\} \mathbf{1}[j \in W(a)] \right], \end{aligned} \tag{2.9}$$

and

$$\begin{aligned} H_j(i,x) &= \Pr\{w(a)=j|x;p_{-i};a_i \neq j\} \\ &= \sum_{a\in A}\left[\frac{1}{|W(a)|}\Pr\{a|x;p_{-i};a_i\neq j\}\mathbf{1}[j\in W(a)]\right]. \end{aligned} \tag{2.10}$$

2.6.3 Equilibrium

We use the notation from Section 2.6.2 to express the probability of victory for candidate j, given a candidate strategy profile x, as a sum over the voters. This sum is given in the following lemma.

Lemma 20 *Given a policy profile, $x \in Y$, the probability of victory by candidate j is given by*

$$R_j(x) = \frac{1}{N}\sum_{i}^{N}[p_i^j(x)G_j(i,x) + (1-p_i^j(x))H_j(i,x)]. \tag{2.11}$$

Proof: Consider any voter i and any candidate j. From the definition of conditional probabilities and the assumption of independence,

$$\begin{aligned} R_j(x) &= \Pr[a_i = j \ \cap \ w(a)=j] + \Pr[a_i \neq j \ \cap \ w(a) = j] \\ &= \Pr[a_i=j]\Pr[w(a)=j|a_i=j] + \Pr[a_i\neq j]\Pr[w(a)=j|a_i\neq j] \\ &= p_i^j G_j(i,x) + (1-p_i^j)H_j(i,x). \end{aligned}$$

The result then follows immediately by summing over i. ∎

Given Equations 2.1 and 2.11, the following lemmata state the first derivatives of each candidate's expected vote and probability of victory with respect to her own strategy, respectively.

Lemma 21 *For any electoral game with differentiable response functions p, any candidate $j \in \mathcal{J}$, and any policy profile $x \in Y$,*

$$D_{x_j} V_j(x) = \sum_{i \in \mathcal{N}} D_{x_j} p_i^j(x). \tag{2.12}$$

Lemma 22 *For any electoral game with differentiable response functions p, any candidate $j \in \mathcal{J}$, and any policy profile $x \in Y$,*

$$D_{x_j} R_j(x) = \sum_{i \in \mathcal{N}} \delta_i^j(p_{-i}(x)) D_{x_j} p_i^j(x) \tag{2.13}$$

where $G_i(j;x)$ and $H_i(j;x)$ are defined in Equations 2.9 and 2.10, respectively, and their difference, $\delta_i^j(p_{-i}(x))$, equals i's pivot probability with respect to candidate j.

Proof:

$$\begin{aligned} R_l(x) &= \sum_{a \in A} \frac{1}{|W(a)|} \mathbf{1}[l \in W(a)] \Pr[a|p(x)] \\ &= \sum_{a \in A: l \in W(a)} \frac{1}{|W(a)|} \prod_{i=1}^{N} p_i^{a_i}(x). \end{aligned}$$

$$\begin{aligned} D_{x_l} R_l(x) &= \sum_{a \in A: l \in W(a)} \left[\frac{1}{|W(a)|} \sum_{i=1}^{N} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \\ &= \sum_{k=1}^{J} \frac{1}{k} \left[\sum_{a \in A: l \in W(a), |W(a)|=k} \left[\sum_{i=1}^{N} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \right] \\ &= \sum_{k=1}^{J} \frac{1}{k} \left[\sum_{i=1}^{N} \left[\sum_{a \in A: l \in W(a), |W(a)|=k} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \right] \quad (2.14) \\ &= \sum_{k=1}^{J} \frac{1}{k} \left[\sum_{i=1}^{N} \left[\sum_{a \in A: l \in W(a), |W(a)|=k, a_i = l} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{l}(x) \right. \right. \\ &\quad \left. \left. + \sum_{a \in A: l \in W(a), |W(a)|=k, a_i \neq l} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \right]. \quad (2.15) \end{aligned}$$

For any voter $i \in \mathcal{N}$ and any vector of policy proposals $x \in Y$, $\sum_{l=1}^{J} p_i^l(x) = 1$, so that, for any candidate $j \in \mathcal{J}$, $\sum_{l=1}^{J} D_j p_i^l(x) = 0$. For any voter i and any candidate j, let $\mathcal{D}(i;j) \subset A$ denote the vote vectors in which i is decisive (or pivotal) for candidate j; i.e., all vote vectors $\hat{a}$ such that $\hat{a}_i = j$, $j \in W(\hat{a})$, and $W(a^{prime})$ and $j \notin W(a'$ for any $a^{prime} = (\hat{a}_1, \ldots, \hat{a}_{i-1}, a_i^{prime}, \hat{a}_{i+1}, \ldots, \hat{a}_N)$ with $a^{prime} \neq j$. Rewriting Equation 2.15:

$$\begin{aligned} D_{x_l} R_j(x) &= \sum_{k=1}^{J} \frac{1}{k} \left[\sum_{i=1}^{N} \left[\sum_{a \in \mathcal{D}(i;l): l \in W(a), |W(a)|=k, a_i = l} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^l(x) \right. \right. \\ &\quad + \sum_{a \notin \mathcal{D}(i;j): l \in W(a), |W(a)|=k, a_i = l,} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^l(x) \\ &\quad \left. \left. + \sum_{a \in A: l \in W(a), |W(a)|=k, a_i \neq l} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \right]. \end{aligned} \tag{2.16}$$

For any voter i, any candidate j, and any vote vector $a \in A$, $a_i \neq j$ implies that $a \notin \mathcal{D}(i;j)$. Thus, it is possible to combine the second and third inner sums in Equation 2.16 and obtain

$$\begin{aligned} D_{x_l} R_j(x) &= \sum_{k=1}^{J} \frac{1}{k} \left[\sum_{i=1}^{N} \left[\sum_{a \in \mathcal{D}(i;l): l \in W(a), |W(a)|=k, a_i = l} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^l(x) \right. \right. \\ &\quad \left. \left. + \sum_{a \notin \mathcal{D}(i;j): l \in W(a), |W(a)|=k} \left[\prod_{j \neq i} p_j^{a_j}(x) \right] D_{x_l} p_i^{a_i}(x) \right] \right]. \end{aligned} \tag{2.17}$$

For any voter $i \in \mathcal{N}$ and candidate $j \in \mathcal{J}$, let $\lceil(i;j) \subset A_{-i}$ denote the set of vectors of votes other than i's in which $j \in W(A)$ and i can not be pivotal for j. That is, regardless of i's vote, $W(a)$ remains the same (and includes j). Formally,

$$\mathcal{D}(i;j) = \{a_{-i} \in A_{-i} : j \in W(a_i; a_{-i}) \forall a_i \in \mathcal{J}\}.$$

Rewriting Equation 2.17,

$$
\begin{aligned}
D_{x_l}R_j(x) &= \sum_{k=1}^{J}\frac{1}{k}\left[\sum_{i=1}^{N}\left[\sum_{a\in\mathcal{D}(i;l):l\in W(a),|W(a)|=k,a_i=l}\left[\prod_{j\neq i}p_j^{a_j}(x)\right]D_{x_l}p_i^l(x)\right.\right. \\
&\quad \left.\left. + \sum_{a_{-i}\in\Gamma(i;j):|W(a)|=k}\left[\sum_{m=1}^{J}\left[\prod_{j\neq i}p_j^{a_j}(x)\right]D_{x_l}p_i^m(x)\right]\right]\right]. \qquad (2.18)
\end{aligned}
$$

Since $\sum_{m=1}^{J}D_jp_i^m(x)=0$ for any $i\in\mathcal{N}$ and $x\in Y$, the second inner sum in Equation 2.18 vanishes, leaving

$$
\begin{aligned}
D_{x_l}R_j(x) &= \sum_{k=1}^{J}\frac{1}{k}\left[\sum_{i=1}^{N}\left[\sum_{a\in\mathcal{D}(i;l):l\in W(a),|W(a)|=k,a_i=l}\left[\prod_{j\neq i}p_j^{a_j}(x)\right]D_{x_l}p_i^l(x)\right]\right] \\
&= \sum_{i=1}^{N}\left[\sum_{a\in\mathcal{D}(i;l)}\frac{1}{|W(a)|}\left[\prod_{j\neq i}p_j^{a_j}(x)\right]D_{x_l}p_i^l(x)\right] \qquad (2.19)
\end{aligned}
$$

The pivot probability of voter i with respect to candidate l, given a policy profile $x\in Y$ and other voters' response functions p_{-i}, is equal to

$$
\delta_i^l(p_{-i}(x)) = \sum_{a\in\mathcal{D}(i;l)}\left[\frac{1}{|W(a)|}\prod_{j\neq i}p_j^{a_j}(x)\right]. \qquad (2.20)
$$

Using Equation 2.20 and substituting δ_i^l into Equation 2.19, we obtain

$$
D_{x_l}R_l(x) = \sum_{i=1}^{N}\delta_i^j(x)D_{x_l}p_i^l(x),
$$

as was to be shown. ∎

We now define p-symmetric, local, interior, and p-interior vectors of response functions. In words, these conditions are fairly straight forward: p-symmetry holds whenever, at some policy profile $x\in Y$, all voters behave identically in expectation; two policy profiles $x,y\in Y$ are local if the distance between them is not too great; a policy profile $x\in Y$ is interior to Y if it is not on the boundary of Y; and a policy profile $x\in Y$ is p-interior if no voter is voting for any candidate with probability zero, *conditional* upon x being the announced policy positions.

Our definition of p-symmetry is satisfied when, conditional on the policy announced, each voter's expected behavior is identical. That is, p-symmetry is a pointwise satisfaction of Assumption 9, as used earlier in Section 2.5.1.

Definition 23 *Given a vector of response functions, p, a policy profile $y \in Y$ is* p-symmetric *if, for all $i, j \in \mathcal{N}$ and all $k \in J$,*

$$p_i^k(y) = p_j^k(y).$$

Any policy profile which is not p-symmetric is referred to as p-asymmetric.

Two policy profiles are ε-local if the distance between them is no greater than ε.

Definition 24 *Two policy profiles, $x, y \in Y$, are ε-local if*

$$||x - y|| \leq \varepsilon.$$

Our methods of proof utilize the necessary first order and sufficient second order conditions for optimization. In order to simplify the analysis, we examine interior equilibria - equilibria for which the necessary first order conditions are that the gradient vanish. Perhaps obviously, an interior equilibrium is any equilibrium in pure strategies, x^*, where x^* is in the interior of Y.

Definition 25 *A vector $x \in Y$ is in the interior of Y, written $x \in Int(Y)$, if, for some $\varepsilon > 0$,*

$$(||z - x|| \leq \varepsilon) \Rightarrow z \in Y.$$

We provide results for p-symmetric local equilibria at p-interior vectors of response functions. The reason for this restriction is that the partial derivatives of a candidate's probability of victory with respect to any voter's response function are functions of the other voters' response functions. Hence, p-interior vectors of response functions *imply that these partial derivatives do not vanish.*

Definition 26 *A vector of response functions p is p-*interior *at a point $x \in Y$ if, for all voters $i \in \mathcal{N}$ and candidates $j \in \mathcal{J}$, $p_i^j(x) > 0$.*

We now define a strict local equilibrium of an electoral game.

Definition 27 *Let $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$ be an electoral game, let $x \in Y$ be a candidate policy profile, and let x'_j be any unilateral deviation by candidate j from x. Then x is a* strict local equilibrium of Γ *if there exists $\varepsilon^* > 0$ satisfying the following. For all $j \in \mathcal{J}$ and for all x'_j which are ε^*-local to x,*

$$u_j(x) \leq u_j(x'_j),$$

with the inequality being strict whenever $x'_j \neq x_j$.

To simplify exposition, whenever we are comparing two electoral games $\Gamma = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u)$ and $\Gamma' = (\mathcal{J}, \mathcal{J}_0, X, \mathcal{N}, p, u')$, we will refer to the set of strict local equilibria of Γ and Γ' as being strict local equilibria under u and strict p-local equilibria under u', respectively.

Finally, for any voter i, any policy profile $x \in Y$, and any given vector of the other voters' response functions, p_{-i}, we will denote the probability of a vector of other voters' actions, $a_{-i} \in A_{-i}$, such that candidate j is either tied for the lead or one vote behind the leading candidate by $\delta_i^j(p_{-i}(x))$. This probability is also referred to as the *pivot probability* and is dealt with quite extensively in Chapter 3.

2.6.4 Sufficient Conditions for Equilibrium Equivalence [5]

In this section, we provide two results which, when taken together, provide an insight into when local equilibria under maximization of expected vote share and probability of victory are equivalent. In essence, p-symmetric local equilibria in which voter behavior is "concave enough" are equivalent. The p-symmetry condition can be linked

[5] **I am extremely grateful to John Duggan for many helpful comments on this topic, including the discovery of a serious error in an earlier draft of this section.**

to the sufficient conditions found in Theorem 13 - voters, at least locally, "look the same" to the candidates. Both of the main results in this section (Theorems 28 and 29) generalize and closely mirror results (Theorems 1 and 4, respectively) in recent work by Duggan (2000).

Duggan (2000) examines the question of equilibrium equivalence between maximization of probability of victory and expected vote share maximization in 2 candidate elections without abstention. In particular, Duggan shows that, in such elections, strict interior Nash equilibria under probability of victory maximization are also equilibria under maximization of expected vote share when the voters' types are independent. As for the converse, Duggan examines a general model of probabilistic behavior known as the *additive bia model*, in which voters' types are represented by a utility bias in favor of one candidate or the other. He proves that the negative definiteness of the matrix of second derivatives of the sum of the voters' utility functions is a sufficient condition for a strict interior Nash equilibrium under maximization of expected vote share to be a Nash equilibrium under maximization of probability of victory.

Our main contributions are to allow for more than 2 candidates in both results and allow for a broader class of voter behavior in Theorem 29.

We now show that, if $x^* \in Y$ is a p-symmetric local equilibrium under either expected vote share or probability of victory maximization, then the negative definiteness of

$$\sum_{i \in \mathcal{N}} D^2 p_i^j(x^*) \tag{2.21}$$

for each candidate j is a sufficient condition for equilibrium equivalence. We also show that this condition can be relaxed to local strict concavity in each candidate j's policy position of

$$\sum_{i \in \mathcal{N}} p_i^j(x^*) \tag{2.22}$$

for local equilibria under probability of victory maximization to also be local equilibria under expected vote share maximization.

Theorem 28 *Let $x^* \in \mathbb{R}^M$ be an interior p-symmetric strict local equilibrium under probability of victory maximization in which $p_i^j(x^*) > 0$ for all voters $i \in \mathcal{N}$ and candidates $j \in \mathcal{J}$ and that, for each candidate $j \in \mathcal{J}$, the following condition is satisfied:*

$$\sum_{i \in \mathcal{N}} p_i^j(x^*) \tag{2.23}$$

is strictly concave with respect to x_j in an open neighborhood of x_j^*. Then x^* is also a strict local equilibrium under maximization of expected vote share.

Proof: Since x^* is an interior local equilibrium under maximization of probability of victory, then the following condition must hold for any candidate $j \in \mathcal{J}$.

$$DR_j(x^*) = \sum_{i \in \mathcal{N}} [\delta_i^j(p_{-i}(x^*)) Dp_i^j(x^*)] = 0.$$

The first order conditions for an interior local equilibrium under maximization of expected vote share are given by the following.

$$DV_j(x^*) = \sum_{i \in \mathcal{N}} Dp_i^j(x^*) = 0.$$

By p-symmetry, it follows that each voter is equally likely to be pivotal, so that $\delta_i^j(p_{-i}(x^*)) = \delta_k^j(p_{-k}(x^*))$ for all $i, k \in \mathcal{N}$ and all $j \in \mathcal{J}$. Therefore, the first order conditions for local equilibrium under expected vote share are satisfied at x^*.

By the hypothesis that $\sum_{i \in \mathcal{N}} p_i^j(x^*)$ is strictly concave in x_j in an open neighborhood of x^*, x_j^* is a strict local maximizer of $V_j(x^*)$ in an open neighborhood of x^* for all candidates j, implying the result. ∎

If the second derivative of expected vote share is negative definite at an interior p-symmetric policy profile x^* for each candidate j, then a strict local equilibrium under maximization of probability of victory is also a strict local equilibrium under expected vote share maximization.

Theorem 29 *Let $x^* \in \mathbb{R}^M$ be an interior p-symmetric strict local equilibrium under expected vote maximization in which $p_i^j(x^*) > 0$ for all voters $i \in \mathcal{N}$ and candidates $j \in \mathcal{J}$. If $D^2V_j(x^*)$ is negative definite for all j, then x^* is also a strict local equilibrium under probability of victory maximization.*

Proof: The first order condition for an interior local equilibrium under expected vote maximization is $DV_j(x^*) = 0$, for all $j \in \mathcal{J}$. To see that this implies the first order condition for local equilibrium under probability of victory maximization, $DR_j(x^*) = 0$ for all $j \in \mathcal{J}$, note that

$$DR_j = \delta^j \cdot Dp^j,$$

where δ^j is the $N \times 1$ vector of pivot probabilities for each voter with respect to candidate j, and Dp^j is the $N \times M$ matrix of partial derivatives of each voter's individual probability of voting for candidate j, and

$$DV_j = Dp^j.$$

Since $\delta > 0$ by the interior nature of p_i for all $i \in \mathcal{N}$ and $\delta_i^j = \delta_k^j$ for all voters $i, k \in \mathcal{N}$ and candidates $j \in \mathcal{J}$, it follows immediately that

$$(DV_j = 0) \Leftrightarrow (Dp^j = 0) \Leftrightarrow (DR_j = 0).$$

A sufficient second order condition for local equilibrium under maximization of R at x^* is the negative definiteness of

$$D^2R_j(x^*) = \sum_{i \in \mathcal{N}} [D\delta_i^j \cdot Dp_i^j + \delta_i^j D^2 p_i^j], \tag{2.24}$$

where $D\delta_i^j = D\delta_k^j$ for all $i, k \in \mathcal{N}$ and $j \in \mathcal{J}$ by the p-symmetry of x^*, so that the first term on the RHS of Equation 2.24 disappears at x^*, since $\sum_{i\in\mathcal{N}} Dp_i^j = Dp^j = 0$. Thus, by the fact that each voter is equally likely to be pivotal,

$$\begin{aligned} D^2R_j(x^*) &= \sum_{i\in\mathcal{N}} \delta_i^j D^2 p_i^j \\ &= \delta_k^j \sum_{i\in\mathcal{N}} D^2 p_i^j, \end{aligned}$$

for any voter $k \in \mathcal{N}$, so that D^2R_j is a positive scalar multiple of $\sum_{i\in\mathcal{N}} D^2p_i^j$, implying that D^2R_j assumes the definiteness of $\sum_{i\in\mathcal{N}} D^2p_i^j$. Finally, since

$$D^2V_j = \sum_{i\in\mathcal{N}} D^2p_i^j,$$

it follows immediately that $D^2R_j(x^*)$ assumes the definiteness of $D^2V_j(x^*)$. By hypothesis, D^2V_j is negative definite, implying the desired result. ∎

An immediate question is what conditions on voter response functions would imply that $D^2V_j(x^*)$ is negative definite? It turns out that a sufficient condition is local concavity of each voter's behavior, coupled with the existence, for each candidate, of at least one voter whose behavior at x^* has a negative definite matrix of second partial derivatives with respect to that candidate's policy choice. The next lemma is straight forward.

Lemma 30 *Suppose that for all $i \in \mathcal{N}$ and some point x^*, p_i is concave in x_j at x^* for each $j \in \mathcal{J}$, and, for each candidate $k \in \mathcal{J}$, there exists at least one voter $l_k \in \mathcal{N}$ for whom $D^2p_{l_k}^k(x^*)$ is negative definite. Then $D^2V_j(x^*)$ is negative definite for each $j \in \mathcal{J}$.*

Proof: Choose any candidate j. We have shown that the matrix of second partial derivatives of v_j is given by

$$D^2V_j(x^*) = \sum_{i\in\mathcal{N}} D^2p_i^j(x^*).$$

The sum of negative semidefinite matrices is negative semidefinite, and the matrix of second partial derivatives of a concave function is a negative semidefinite matrix. In addition, the sum of any nonnegative number of negative semidefinite matrices and any positive number of negative definite matrices is a negative definite matrix. The result follows immediately by noting that the choice of j is arbitrary. ∎

2.6.5 Necessary Conditions

In this section, we investigate necessary conditions for equilibrium equivalence and best response equivalence between maximization of expected vote share and maximization of probability of victory. We examine necessary conditions for equilibrium equivalence first, as equilibrium equivalence is itself a necessary condition for best response equivalence (Theorem 6 and Corollary 7).

Preliminaries

In this section, we restrict attention to the case where X is a compact subset of a finite dimensional Euclidean space, $\mathbb{R}^M$, where M is an arbitrary integer. We first provide a technical result regarding a representation of voter behavior in our model. In particular, we characterize voter behavior as a point in finite dimensional Euclidean space. We also retain the restriction that each p_i is continuously differentiable on Y.

Given that we are concerned with satisfaction of first order conditions at some vector of proposals $x \in Y$, each continuously differentiable response function p_i is completely characterized by a vector $q_i(x) \in \Delta(\mathcal{J}) \times \mathbb{R}^J$. For a given $x \in Y$, the space of all such $q_i(x)$ is denoted by $Q_i(x) \subset \mathbb{R}^{J(M+1)}$.

Formally, we define $q_i(x)$ as follows:

$$\begin{aligned} q_i(x) &= (p_i(x), Dp_i^1(x), \ldots, Dp_i^J(x)) \\ &= (p_i^1(x), \ldots, p_i^J(x), \frac{\partial p_i^1(x)}{\partial x_1^1}, \ldots, \frac{\partial p_i^1(x)}{\partial x_1^M}, \frac{\partial p_i^2(x)}{\partial x_2^1}, \ldots, \frac{\partial p_i^J(x)}{\partial x_J^M}). \end{aligned}$$

That is, $q_i(x)$ is the concatenation of $p_i(x)$ and its derivatives with respect to each candidate's policy proposal evaluated at x. This definition of q_i turns out to be a sufficient parameterization of the space of continuously differentiable response functions to show that the first order conditions for equilibrium equivalence between maximization of expected plurality and probability of victory are generically not satisfied by vectors of continuously differentiable response functions. Given a vector of response functions $p = \{p_1, \ldots, p_N\}$, we write $q(x) = \{q_1(x), \ldots, q_N(x)\}$.

The next proposition states that $q(x)$ is a sufficient statistic for the first derivatives of both expected vote share and probability of victory with respect to any candidate's policy choice.

Proposition 31 *Choose $x^* \in Int(Y)$ and let p and $\hat{p}$ be vectors of continuously differentiable response functions such that $q(x^*) = \hat{q}(x^*)$. Then*

- $DR_j(p(x^*)) = DR_j(\hat{p}(x^*))$ *and*
- $DV_j(p(x^*)) = DV_j(\hat{p}(x^*))$

for all $j \in \mathcal{J}$.

Similarly, it follows immediately that if two sequences, $\{q(x)_\alpha\}_{\alpha=1}^{\infty}$ and $\{\hat{q}(x)_\alpha\}_{\alpha=1}^{\infty}$, converge, then the sequences of evaluations of each candidate's objective functions converge as well, since V_j and R_j are both continuous functions of q.

Generic Failure of Equilibrium Equivalence

The following results hinge on the fact that the first derivative of a candidate's expected vote share with respect to her own policy choice is a mapping from $\mathbb{R}^N$ to $\mathbb{R}$ while the first derivative of her probability of victory with respect to her policy choice is a mapping from $\mathbb{R}^{2N}$ to $\mathbb{R}$.

For the remainder of this section we restrict attention to the case where the response functions are each continuously differentiable in X and show that the set of continuously differentiable response functions under which expected plurality and probability of victory maximization are equivalent is "small" in a precise sense. In particular, we show that the set of continuously differentiable response functions which yield a given vector q is *shy* in the space of continuously differentiable response functions. First, however, we define a shy set.

Definition 32 *Let X be a complete metrizable topological vector space. A Borel set $E \subset X$ is* shy *if there exists a regular Borel probability measure μ on X with compact support such that $\mu(E + x) = 0$ for every $x \in X$. A (not necessarily Borel) subset $F \subset X$ is shy if it is contained in a shy Borel set. A subset $Y \subset X$ is* prevalent *if $X \setminus Y$ is shy.*

These definitions of shyness and prevalence are due to Hunt, Sauer, and Yorke (1992). The notion of shyness has been generalized by Anderson and Zame (2000) to include a notion of relative shyness. For our purposes, for example, the set of continuously differentiable response functions turns out to be shy (since the ambient vector space is the space of continuous response functions). Thus, we use the following definition of shyness with respect to another set.

Definition 33 *Let X be a topological vector space and let $C \subset X$ be a convex Borel subset of X which is completely metrizable in the relative topology. Fix $c \in C$. A Borel subset $E \subset X$ is* shy in C at c *if for each $\delta > 0$ and each neighborhood W of 0 in X, there is a regular Borel probability measure μ on X with compact support such that supp $\mu \subset [\delta(C - c) + c] \cap (W + c)$ and $\mu(E + x) = 0$ for every $x \in X$. A subset E is* shy in C *if it is shy at each point $c \in C$. A (not necessarily Borel) subset $F \subset C$ is shy in C if it is contained in a shy Borel set. A subset $Y \subset C$ is* prevalent in C *if its complement $C \setminus Y$ is shy in C.*

A stronger version of shyness is *finite shyness*, as defined in Anderson and Zame (2000). For any finite dimensional subspace $V \subset X$, we write λ_V for Lebesgue measure on V.

Definition 34 *A Borel subset $E \subset C$ is* finitely shy in C *if there is a finite-dimensional subspace $V \subset X$ such that $\lambda_V(E + x) = 0$ for every $x \in X$. A (not necessarily Borel) subset $F \subset X$ is finitely shy in C if it is contained in a finitely shy Borel set.*

The notion of shyness extends measure-theoretic notions of genericity to infinite dimensional spaces. Since the space of continuous response functions is infinite dimensional whenever X is infinite, we use shyness as our analogue for "smallness" in this space.

We denote the space of continuously differentiable response functions from Y to $\Delta(\mathcal{J})$ by $\mathcal{P}_{dif}(Y, \mathcal{J})$. This set is a closed subset of the space of continuous functions $\mathcal{C}(Y, \mathbb{R}^J)$. We denote the space of N-dimensional vectors of continuously differentiable response functions by $\mathcal{P}^N_{dif}(Y, \mathcal{J})$.

Our next result states that, for any vector of policy proposals $x \in Y$ any point in the J-dimensional simplex, t_0, and any J-dimensional vector of M-dimensional real vectors $t_1, \ldots, t_J$, the set of continuously differentiable response functions p for which $q = t$ (i.e., $p(x) = t_0$ and $Dp^j(x) = t_j$) is finitely shy in the space of all continuously differentiable response functions. In words, the space of functions which satisfy a given first-order condition is "small."

Theorem 35 *Let Y be compact. Then, given any $\tilde{x} \in Y$ and any $t \in \mathbb{R}^{J+JM}$, the set*

$$C(t) = \{p : q(\tilde{x}) = t\}$$

is finitely shy in $\mathcal{P}_{dif}(Y, \mathcal{J})$, the space of continuously differentiable response functions on Y.

Proof: We first show that $C(t)$ is a Borel subset of $\mathcal{P}_{dif}(Y \mathcal{J})$.

Choose any countable dense subset $\hat{X} \subset Int(X)$. By the continuous differentiability of p, if $p(\tilde{x}) = k$ and $Dp(\tilde{x}) = L$, where $k \in \Delta(\mathcal{J})$ and L is a $J \times J$ real-valued matrix,

$$\lim_{x \to \tilde{x}} \frac{|p(x) - p(\tilde{x}) - L \cdot (x - \tilde{x})|}{||x - \tilde{x}||} = 0.$$

Equivalently, for each $\alpha \in \mathbb{Z}_{++}$, there exists $\beta \in \mathbb{Z}_{++}$ such that

$$\frac{|p(x) - p(\tilde{x}) - L \cdot (\hat{x} - \tilde{x})|}{||\hat{x} - \tilde{x}||} < \frac{1}{\alpha}$$

for any $\hat{x} \in \hat{X}$ such that $||\hat{x} - x^*|| < \frac{1}{\beta}$.

For any $\alpha, \beta \in \mathbb{Z}_{++}$ and $\hat{x} \in \hat{X}$,

$$C(t; \alpha, \beta, \hat{x}) = \left\{ p \in \mathcal{P}_{dif}(Y, \mathcal{J}) : \frac{|p(\hat{x}) - p(\tilde{x}) - L \cdot (\hat{x} - \tilde{x})|}{||\hat{x} - \tilde{x}||} < \frac{1}{\alpha} \right\},$$

which is easily verified to be an open subset of $\mathcal{P}_{dif}(Y, \mathcal{J})$.

Now let

$$C(t) = \cap_{\alpha=1}^{\infty} \cup_{\beta=1}^{\infty} \left((\cap_{0<||\hat{x}-x^*||<\frac{1}{\beta}} C(t; \alpha, \beta, \hat{x}) \right),$$

which, as the countable union and intersection of open subsets of $\mathcal{P}_{dif}(Y, \mathcal{J})$, is a Borel subset of $\mathcal{P}_{dif}(Y, \mathcal{J})$.

We now show that $C(t)$ is finitely shy in $\mathcal{P}_{dif}(Y, \mathcal{J})$. In order to see this, choose $x^* \in Y$ and define

$$h_{x^*}(x; \phi) = (\frac{e^{\phi||x_1-x^*||}}{2(e^{\phi||x_1-x^*||} + e^{-t||x_1-x^*||})}, \frac{e^{-\phi||x_1-x^*||}}{2(e^{\phi||x_1-x^*||} + e^{-\phi||x_1-x^*||})}, \frac{1}{2(J-2)}, \ldots, \frac{1}{2(J-2)}). \tag{2.25}$$

It is simple to verify that $h_{x^*}(\cdot; \phi) \in \mathcal{P}_{dif}(Y, \mathcal{J})$ for all $x^* \in Int(Y)$ and all $\phi \in \mathbb{R}$.

Let

$$H = \{h_{x^*}(x; \phi) : \phi \in \mathbb{R}\}$$

denote the one-dimensional subspace of $\mathcal{C}(Y; \mathbb{R}^J)$ which is spanned by $h_{x^*}(x; \phi)$. Since every element of H is also an element of $\mathcal{P}_{dif}(Y, \mathcal{J})$, it follows that $\lambda_H(\mathcal{P}_{dif}(Y, \mathcal{J})) > 0$. Now we must check that $\lambda_H(C(t) + g) = 0$ for any $g \in \mathcal{C}(Y, \mathbb{R}^J)$. It suffices to show that, for any $g \in \mathcal{C}(Y, \mathbb{R}^J)$, $(H - g) \cap C(t)$ is empty or a singleton. Suppose, by way of contradiction, that this were not the case. Then there would exist some $g \in \mathcal{C}(Y, \mathbb{R}^J)$, some distinct $\phi_1, \phi_2 \in \mathbb{R}$ and some distinct $p_1, p_2 \in C(t)$ such that

$$\begin{aligned} h(\cdot;\phi_1) - g &= p_1 \\ h(\cdot;\phi_2) - g &= p_2. \end{aligned}$$

If this were the case, then it would follow that

$$h(\cdot;\phi_1) - h(\cdot;\phi_2) = p_1 - p_2. \tag{2.26}$$

Since, by construction, all four functions in Equation 2.26 are everywhere continuously differentiable in Y, we can take the first derivative of the first component of each of these vector-valued functions:

$$\begin{aligned} D(h^1(x^*;\phi_1) - h^1(x^*;\phi_2)) &= D(p_1^1(x^*) - p_2^1(x^*)) \\ \phi_1 - \phi_2 &= Dp_1 - Dp_2 \\ &= q_1^1 - q_2^1 \\ &= 0, \end{aligned}$$

contradicting the supposition that $\phi_1 \neq \phi_2$. Thus, $(H - g) \cap C(t)$ is either empty or a singleton, implying that $\lambda_H(C(t) + g) = 0$ for every $g \in \mathcal{C}(Y, \mathbb{R}^J)$. Thus, $C(t)$ is finitely shy in $\mathcal{P}_{dif}(Y, \mathcal{J})$. Since t is arbitrary, the result follows. ∎

Theorem 35 states that the set of continuously differentiable response functions which possess a given value of $q(\tilde{x})$ at a given point in $\tilde{x} \in Y$ is finitely shy in (i.e., small relative to) the space of all continuously differentiable response functions. The next result strengthens this finding by showing that the set of continuously differentiable response functions which, for a given point $\tilde{x} \in Y$, possess a given value, $p(\tilde{x}) = \tilde{p} \in \Delta(\mathcal{J})$, and vector of derivatives, $Dp(\tilde{x}) = \tilde{d} = \{Dp^1(\tilde{x}), \ldots, Dp^J(\tilde{x})\}$, (i.e., all continuously differentiable p for which $q(\tilde{x}) = (\tilde{p}, \tilde{d})$) is finitely shy in the space of all continuously differentiable response functions possessing a derivative equal to $\tilde{d}$ (i.e., all continuously differentiable p for which $q(\tilde{x}) = (t, \tilde{d})$ for some $t \in \Delta(\mathcal{J})$.

Theorem 36 *Given any $\tilde{x} \in Y$, any $\tilde{p} \in \Delta(J)$, and any $\tilde{d} \in \mathbb{R}^J$, the set*

$$C(\tilde{p}, \tilde{d}) = \{p \in \mathcal{P}_{dif}(Y, \mathcal{J}) : q(\tilde{x}) = (\tilde{p}, \tilde{d})\}$$

is finitely shy in

$$B(\tilde{d}) = \{p \in \mathcal{P}_{dif}(Y, \mathcal{J}) : Dp(\tilde{x}) = \tilde{d}\}.$$

Proof: We show that $C(\tilde{p}, \tilde{d})$ is a Borel subset of $B(\tilde{d})$. Define

$$\bar{C}(t; \tilde{d}) = \{p \in \mathcal{P}_{dif}(Y, \mathcal{J}) : q(\tilde{x}) = (s, \tilde{d}), s > t\},$$

and

$$\underline{C}(t; \tilde{d}) = \{p \in \mathcal{P}_{dif}(Y, \mathcal{J}) : q(\tilde{x}) = (s, \tilde{d}), s < t\},$$

where $>$ denotes the usual partial ordering of $\mathbb{R}^J$. It is simple to see that $\bar{C}(t; \tilde{d})$ is an open subset of $B(\tilde{d})$. Denoting the set of J-dimensional vectors of rational numbers by $\mathbb{Q}$, noting that the set of J-dimensional rationals are a countably dense subset of $\mathbb{R}^J$, and observing that

$$C(\tilde{p}, \tilde{d}) = \left[\cap_{t>\tilde{p}, t\in\mathbb{Q}}[\underline{C}(t; \tilde{p})]\right] \cap \left[\cap_{t<\tilde{p}, t\in\mathbb{Q}}[\bar{C}(t; \tilde{p})]\right],$$

we see that $C(\tilde{p}, \tilde{d})$ is a countable intersection of open subsets of $B(\tilde{d})$. Hence, $C(\tilde{p}, \tilde{d})$ is a Borel subset of $B(\tilde{d})$.

Now choose $r = \{r_1, \ldots, r_N\} \in (\mathcal{P}_{dif}(Y, \mathcal{J}))^N$ satisfying

$$\begin{aligned} r_i &= \tilde{p}_i \text{ and} \\ Dr_i &= \tilde{d}_i \end{aligned}$$

for all $i \in \{1, \ldots, N\}$. Obviously, $r \in B(\tilde{p})$ and $r \in C(\tilde{p}, \tilde{d})$. Next, define

$$h(r;t) = (r_1 + t, \ldots, r_N + t), \tag{2.27}$$

for all $t \in \mathbb{R}^J$ that sum to zero. That is,

$$t \in T = \{t \in \mathbb{R}^J : \sum_{i=1}^{J} t_i = 0\}.$$

Let $H(r)$ denote the one-dimensional subspace of $\mathcal{C}(Y, \mathbb{R}^J)$ spanned by $h(r;t)$ and $\lambda_{H(r)}$ denote Lebesgue measure relative to $H(r)$

Recall that

- $p_i(x) \in (0,1)^J$ for all $i \in \mathcal{N}$ and all $x \in Y$,
- p_i is continuously differentiable for all $i \in \mathcal{N}$,
- Y is compact, and
- both $\mathcal{J}$ and $\mathcal{N}$ are finite sets.

Thus, there must exist $\varepsilon^* > 0$ satisfying the following:

$$\varepsilon^* = \min_{i \in \mathcal{N}} \left[\min_{j \in \mathcal{J}} \left[\min_{x \in Y} [p_i^j(x)] \right] \right] > 0.$$

For all t in the following set:

$$\bar{T} = \{t \in \mathbb{R}^J : \sum_{i=1}^{J} t_i = 0, \max_{1 \le i \le J} t_i < \varepsilon^*\},$$

it follows that, as defined in Equation 2.27, $h(r;t) \in \mathcal{P}_{dif}(Y, \mathcal{J})$ for all $t \in \bar{T}$. Furthermore, $Dh_i(r;t) = \tilde{d}_i$ for all $t \in T$, so that $h(r;t) \in B(\tilde{d})$ for all $t \in \bar{T}$. Since $\bar{T}$ is a nonempty open subset of T, restricting t to $\bar{T}$ determines an open subset of $H(r)$, implying that $\lambda_{H(r)}(B(\tilde{d})) > 0$.

For simplicity, let $C = C(\tilde{p}, \tilde{d})$ and $H = H(r)$. Now we claim that $\lambda_H(C + g) = 0$

for any $g\mathcal{C}(Y;\mathbb{R}^J)$. To see this, it suffices to show that

$$(H-g)\cap C$$

is empty or a singleton.

Suppose, by way of contradiction, that $(H-g)\cap C$ has at least two elements for some $g\in\mathcal{C}(Y,\mathbb{R}^J)$. Then, for some $g\in\mathcal{C}(Y,\mathbb{R}^J)$, there must exist distinct $s,t\in T$ and distinct $a,b\in C$ such that

$$\begin{aligned} r+s-g &= a \\ r+t-g &= b. \end{aligned}$$

This would imply that

$$\begin{aligned} s-t &= a-b, \\ s-t &= 0, \end{aligned}$$

implying that s and t are not distinct, resulting in a contradiction. (Recall that $s,t\in\mathbb{R}^J$, so that $s-t=0$ implies that $s_i=t_i$ for all $i\in\mathcal{J}$.) Since the supposition that $(H-g)\cap C$ has more than one element results in a contradiction, it must be the case that $(H-g)\cap C$ has at most one element, implying that $\lambda_H(C)=0$, so that $C(\tilde{p},\tilde{d})$ is finitely shy in $B(\tilde{d})$. ∎

Theorems 35 and 36 together imply our main negative result concerning equilibrium equivalence between maximization of expected vote share and probability of victory maximization: for any finite electorate, the set of continuously differentiable response functions which simultaneously satisfy the necessary first order conditions for local equilibrium equivalence is finitely shy in the set of continuously differentiable response functions which satisfy the necessary first order conditions for expected vote share maximization.

In particular, for any finite positive number of voters N, any finite positive number of candidates J, and any point $x^* \in Int(Y)$, let $P_V^N(x^*, J) \subset \mathcal{P}_{dif}^N(Y, \mathcal{J})$ denote the set of N-dimensional vectors of continuously differentiable response functions such that x^* satisfies the necessary first order conditions for local equilibrium under maximization of expected vote share, V, and let $P_{V,R}^N(x^*, J) \subset P_V^N(x^*, J)$ denote the set of N-dimensional vectors of continuously differentiable response functions which satisfy the necessary first order conditions for local equilibrium under both expected vote share maximization and maximization of probability of victory, R, at x^*. The next theorem states that $P_{V,R}^N(x^*, J)$ is finitely shy in $P_V^N(x^*, J)$ for any finite N and any $x^* \in Int(Y)$.

Formally, for finite and positive N and J, compact X, and any $x^* \in Int(Y)$, $P_V^N(x^*, J)$ is defined as any $p \in \mathcal{P}_{dif}^N(Y, \mathcal{J})$ for which, for all $j \in \{1, \ldots, N\}$,

$$DV_j(x^*) = \sum_{i=1}^{N} Dp_i^j(x^*) = 0,$$

and $P_R^N(x^*, J)$ is defined as any $p \in \mathcal{P}_{dif}^N(Y, \mathcal{J})$ for which, for all $j \in \{1, \ldots, N\}$,

$$DR_j(x^*) = \sum_{i=1}^{N} \delta_i^j Dp_i^j(x^*) = 0.$$

Finally, $P_{V,R}^N(x^*, J) = P_V^N(x^*, J) \cap P_R^N(x^*, J)$.

Theorem 37 *For any finite, positive N, any finite, positive J, any compact policy space X, and any point $x^* \in Int(Y)$, $P_{V,R}^N(x^*, J)$ is finitely shy in $P_V^N(x^*, J)$.*

Proof: Fix $N, J \in \mathbb{Z}_{++}$, choose any compact $X \subset \mathbb{R}^M$, and choose any $x^* \in Int(Y)$. Now select any $p \in P_V^N(x^*, J)$. We show that any collection of $N-1$ voter response functions exactly pins down the necessary $q_i(x^*)$ for the final voter i. We consider voter N without loss of generality.

Note that $Dp_N^j(x^*)$ is uniquely determined by $Dp_1^j(x^*), \ldots, Dp_{N-1}^j(x^*)$ because, since $p \in P_V^N(x^*, J)$, it must be the case that, for each $j \in \{1, \ldots, J\}$,

$$\sum_{i \in \mathcal{N}} Dp_i^j(x^*) = 0.$$

Furthermore, note that $p \in P^N_{V,R}(x^*, J)$ implies that, for each $j \in \{1, \ldots, J\}$,

$$\sum_{i \in \mathcal{N}} \delta^j_i(p_{-i}(x^*)) Dp^j_i(x^*) = 0,$$

implying that the N^{th} individual's probability of voting for each candidate j, $p^j_N(x^*)$, is uniquely determined by $p_1, \ldots, p_{N-1}$ and $Dp^j_1(x^*), \ldots, Dp^j_N(x^*)$.

Thus, given p_{-N}, $Dp^j_N(x^*)$ is uniquely determined for each $j \in \mathcal{J}$ by the fact that x^* satisfies each candidate's necessary first order conditions for maximization of expected vote share. Furthermore, this uniquely determines $p^j_N(x^*)$ for each candidate $j \in \mathcal{J}$. Denote the determined value of $Dp^j_N(x^*)$ by $\tilde{d}_{-N}$, and the determined value of $p^j_N(x^*)$ by $\tilde{p}_{-N}$, where the subscripts emphasize the manner in which these values are determined.

Applying Theorem 36, it follows that the set of continuously differentiable response functions which satisfy both $Dp^j_N(x^*) = \tilde{d}_{-N}$ and $p^j_N(x^*) = \tilde{p}_{-N}$ is finitely shy in the set of continuously differentiable response functions which satisfy $Dp^j_N(x^*) = \tilde{d}_{-N}$. The result follows by noting that the ordering of voters is arbitrary. ∎

Theorem 37 states that interior equilibrium equivalence is an incredibly rare event with respect to distributions of voter behavior. Indeed, the result is stronger: "critical point" equivalence is a rare event. This result may seem alternatively surprising or expected, depending on how one views our parameterization of the space of response functions. One requirement we make is that voter behavior be p-interior and continuously differentiable in x - this rules out deterministic voting models, for example. We also do not require that behavior satisfy any normative properties, such as symmetry - one might think that $p^1_i(x, y)$ should equal $p^2_i(y, x)$, for example. Investigation of further restrictions on voter behavior is an important research topic for future work. In particular, if voter response functions are symmetric, then the zero sum nature of candidate competition in the absence of abstention will lead to no candidate receiving as probability of victory, or an expected vote share, of less than $\frac{1}{J}$ in a pure strategy equilibrium. A final point is that we do not deal with mixed strategy equilibrium equivalence. An investigation of this question might be a worthwhile topic for future research as well.

Generic Failure of Best Response Equivalence

The next result, which follows immediately from Theorem 37, states that, so long as voter behavior is continuously differentiable with respect to the policy positions of the candidates and at least one candidate's best response correspondence takes at least one value in the interior of X,[6] best rēsponse equivalence between maximization of expected vote and probability of victory maximization will generically not hold in probabilistic voting models.

Theorem 38 *Let there exist at least one candidate $j \in \mathcal{J}$ and one vector of proposals by all candidates other than j, $\hat{x}_{-j} \in Y_{-j}$, such that* $\arg\max_{x \in X} V_j(x_j; \hat{x}_{-j}) \cap Int(X) \neq \emptyset$. *Then, for any finite integer n, any finite set of candidates $\mathcal{J}$ and any compact policy space X, the set of n-dimensional vectors of continuously differentiable response functions for which, for all $j \in \mathcal{J}$ and all $x_{-j} \in Y_{-j}$,*

$$\arg\max_{x_j \in X} V_j(x_j; x_{-j}, p) = \arg\max_{x_j \in X} R_j(x_j; x_{-j}, p)$$

is finitely shy in the space of all n-dimensional vectors of continuously differentiable response functions.

Proof: Consider a candidate j and a vector opponent's proposals $\hat{x}_{-j}$ for which there exists $\hat{x}_j \in Int(X)$ such that

$$DV_j(\hat{x}_j; \hat{x}_{-j}) = 0. \tag{2.28}$$

(The existence of such a candidate and vector of proposals is guaranteed by hypothesis.) By Theorem 37, it follows that the set of continuously differentiable voter

[6]This condition seems mild, given the zero-sum nature of the different objective functions, but we have not explored its implications to any extent.

response functions under which

$$DR_j(\hat{x}_j; \hat{x}_{-j}) = 0 \tag{2.29}$$

is finitely shy in the set of all continuously differentiable voter response functions satisfying Equation 2.28. Since p_i is continuously differentiable on Y for all $i \in \mathcal{N}$ and R_j is differentiable in p by Lemma 11, Equation 2.29 is a necessary condition for $\hat{x}_j$ to be a best response to $\hat{x}_{-j}$. The result follows immediately. ∎

Thus, by supposing that a point is a best response under expected vote share and assuming that voter behavior is continuously differentiable in the policy proposals, we have shown that the set of voter response functions which satisfies the necessary first order conditions for best response equivalence between maximization of probability of victory and expected vote share maximization is small in relation to the entire set of continuous voter response functions. It should be noted that this qualification is a key point. Our arguments are taking any continuous voter response function as a possible form of probabilistic behavior. This assumption may or may not be justified in different areas of inquiry. For example, one may want to restrict attention to those voter response functions which are derivable from a traditional logit model of voter behavior with underlying single-peaked policy preferences (such response functions must intersect the centroid, for example).

2.7 Extensions

There are several questions regarding candidates' objective functions which remain open. A few of these questions are what are the effects of different electoral institutions on equivalence between candidate objective functions, what is the asymptotic behavior of candidate objective functions, what happens when candidates are uncertain about their opponents' objective functions, and how are the behaviors of the voters affected by the objective functions of the candidates?

2.7.1 Electoral Rules

Regarding electoral rules, this chapter has ignored the possibility of proportional representation, multiple winners, multiple ballot systems (e.g., simple majority rule systems with runoffs or party based systems with primaries), and different scoring rules such as approval voting and the Borda count. A positive result which is not particularly surprising is that maximizing probability of victory in a pure (i.e., one without a minimum vote threshold required for representation) proportional representation system is equiavent to maximizing expected vote share. Investigating how the optimal strategies under different candidate objectives change as the electoral system is changed is not only a very interesting topic, but also seemingly tractable.

2.7.2 Asymptotic Equivalence

The issue of asymptotic equivalence has been broached earlier in this chapter in reference to the works of Hinich (1977) and Ledyard (1984). We have shown in two examples (Examples 19 and 18), that asymptotic best response equivalence may not hold in electoral competition. Further, both examples can be shown, in a straight forward fashion, to be robust in the sense that the parameters of each may be perturbed and retain the failure of best response equivalence. However, we have not provided any general results about asymptotic equivalence of candidate objectives. We have not done so for several reasons, of which at least two should be noted.

First, the question of asymptotic equivalence is muddied by at least one significant issue: what exactly does one mean by asymptotic? In what way does one assume that the electorate grow larger? Secondly, it is not clear what the notion of asymptotic equivalence actually means.

Both of these issues lead to a number of somewhat promising routes for future work. With respect to the first issue (how does one assume that larger electorates are generated), one might examine the behavior of the three payoff functions when a given electorate is replicated without bound, or when each voter's response function is drawn independently according to some distribution on the space of possible re-

sponse functions. The positive results presented in this chapter regarding equilibrium equivalence may be used as leverage in such a research project. With respect to the second issue, one might consider several definitions of asymptotic equivalence. A few examples include (1) the existence of a finite number such that, for all electorates with more than this number of voters, one or more of the notions of equivalence defined above holds, (2) best response functions converge asymptotically, or (3) some subset of the equilibria (if they exist) under two objective functions in the sequence of games with finite electorates converge asymptotically.

2.7.3 Strategic Importance of Objective Functions

If we as researchers are uncertain about the true motives of candidates, then it seems reasonable to suppose that at least some candidates are unsure as to their opponents' true motivations. Is such uncertainty important? For example, will different equilibria appear in electoral games in which candidates must account for the fact that their opponent may not care about winning, *per se*, but rather attempt to maximize his or her vote total, conditional upon victory? A preliminary intuition is "probably," as we have provided several examples of situations in which the incentives of candidates differ considerably under maximization of expected vote share and maximization of probability of victory. This question of electoral competition with incomplete information about opponent's payoff functions is a very promising avenue for future research.

2.7.4 Voter Behavior and Preferences

Finally, an important question concerns the behavior of voters. In particular, the incentives of voters are usually assumed to be with respect to the implemented policy. We have assumed that voter behavior is taken as given by the candidates and is invariant to the candidates' preferences. We do this for two reasons. The first reason is methodological - one of the motivations for the study of equivalence is analytical - if objective functions are equivalent, then the analysis of a model under the assumption *of one objective* function is, in the appropriate sense, sufficient analysis of the same model with the equivalent objective function. The second reason is that in many cases it is assumed that voters' preferences are defined over outcomes, and therefore, after the policy announcements are made, the voters are playing a subgame amongst themselves - one in which the candidates have no further role. In some sense, the policy announcements are assumed to have been handed down from outer space. While this is a pessimistic view of the role voters play in the campaign, it is analytically tractable as well as serving as a useful benchmark case. This is not to say that future research should not examine the implications of voter behavior and preferences which are not invariant to candidate objective functions. Indeed, if the candidate incentives under two different candidate objective functions are not identical, then the incentives and abilities of voters to affect candidate behavior under each of the candidate objective functions may differ.[7]

2.8 Conclusions

In this chapter, we have attempted to make several contributions to the formal theory of elections. The first of these is to point out that a rigorous statement and proof of Hinich's (1977) claim that, asymptotically, maximizing plurality and maximizing probability of victory yield equivalent strategies in equilibrium in two candidate elections without voter abstention is not as obvious as might have been assumed. This is important if only because the claim has been widely cited in the literature. We also provide a counterexample to the claim in order to show the need for further investigation into the topic.

The second contribution concerns two candidate elections. It is shown in Theorem 13 that, regardless of the number of voters, maximization of plurality and maximization of probability of victory are *equivalent* objective functions (i.e., they yield identical best response correspondences) in two candidate elections without abstention when voters' behavior satisfies Assumptions 1 and 9.

[7] I thank Richard McKelvey for pointing out this issue.

As stated earlier, Theorem 13 is in some respects weaker, and in others stronger, than Hinich's original statement. Hinich's claim does not require our symmetry condition, Assumption 9. On the other hand, Hinich's claim is asymptotic, while Theorem 13 states that the best response functions are identical for any number of voters.

We have also provided sufficient conditions for local equilibrium equivalence between expected vote share maximization and maximization of probability of victory. By extending arguments due to Duggan (2000), we have shown that "concave enough" aggregate voter behavior is a sufficient condition for local equilibrium equivalence between these objective functions. Conversely, we have shown that the set of continuously differentiable voter response functions which exhibit local equilibrium equivalence is a "small" set of continuously differentiable response functions. That is, local equilibrium equivalence is nongeneric. In particular, we have shown that the set of continuously differentiable response functions which lead, at a given vector of proposals, to local equilibria under both objective functions is small relative to the set of continuously differentiable response functions which lead to a local equilibrium at that same point under maximization of expected vote share.

Finally, we have shown that, generically, best response equivalence does not hold between maximization of expected vote share and maximization of probability of victory in single member, simple plurality elections without abstention. This follows from our genericity result regarding equilibrium equivalence. In particular, if local equilibrium equivalence does not hold, then best response equivalence does not hold.

Chapter 3 Voting in Large Elections [1]

3.1 Introduction

This chapter investigates properties of Quantal Response Equilibrium (see McKelvey and Palfrey (1995), (1998)) in spatial voting games. The Quantal Response Equilibrium (QRE) is a theory of behavior in games that assumes that individuals get privately observed random payoff disturbances for each action available to them. The QRE is then just the Bayesian equilibrium of this game of incomplete information. In a QRE, although voters adopt pure strategies, from the point of view of an outside observer who does not know the payoff disturbance, the players choose between strategies probabilistically, choosing actions that yield higher utility with higher probability than actions that yield lower utility. The probability that one action is chosen over another is based on the utility difference between the alternatives.

In this chapter, we work in a Bayesian framework, as in Ledyard (1984), and take into account the game theoretic considerations for the voters, but unlike Ledyard, we assume that voters have privately observed payoff disturbances associated with each action. Our only restrictions on preferences are that they are uniformly bounded. Further, we consider multi candidate contests. But our results basically extend those of the earlier literature. We find that for large enough electorates there is a convergent equilibrium at the alternative that maximizes social welfare. For two candidate contests, the equilibrium is unique. Our equilibrium is global, as in Lin, Enelow, and Dorussen (1999), but in our model, the conditions for a global equilibrium are satisfied by allowing the number of voters to grow large rather than by assuming the utility shock becomes large.

The main contribution of this research over the previous work is to obtain a global

[1]**This chapter is jointly authored with Richard McKelvey, who acknowledges the financial support of NSF grant #SBR-9631627 to the California Institute of Technology.**

candidate equilibrium in large electorates with very little in the way of assumptions about voter preferences. The main difference between our approach and previous work on probabilistic voting is the way in which we model the probabilistic voting. As in (1984), by treating the voter decisions as a game, we explicitly include the pivot probability in the voters' expected utility calculations. In large electorates, because the probability of being pivotal goes to zero, the expected utility difference between any two candidates also goes to zero. Thus, under the QRE assumptions, the voter's choice is determined mainly by the candidate specific payoff disturbance. Hence, in aggregate, voters vote less based on policy, and more based on candidate attributes as the size of the electorate grows. However, even though individuals become less responsive to policy differences, in large electorates, since the total number of voters is also getting large, there is still enough policy voting at the aggregate level to force the candidates to the social optimum.

3.2 The Model

We assume the existence of a finite dimensional *policy space*, $X \subseteq \Re^m$, where X is bounded, and finite sets N and K of *voters* and *candidates*, respectively. Write $n = |N|$ and $k = |K|$ for the total number of each. We let 0 indicate *abstention*, and write $K_0 = K \cup \{0\}$ for the set of candidates plus abstention.

We assume that for each voter, $i \in N$, there is a space T_i of possible characteristics, or *types* of the voter. Write $T = \Pi_{i \in N} T_i$. We assume that $T_i = \mathcal{T} \times \left(\Re^{K_0}\right)^{\mathcal{T}}$ is partitioned into two parts, representing the policy and consumption based parts, respectively, and that $\mathcal{T}$ is a complete separable metric space. Voters' preferences over the policy space are described by a *utility function*, $u : X \times \mathcal{T} \to \Re$. Hence, the utility of voter $i \in N$, of type $t_i = (\tau_i, \eta_i(\tau_i)) \in T_i$ for the policy $x \in X$ is $u(x, \tau_i)$. Assume that the distribution of the voter i's types is given by an atom-less probability measure of full support, ρ_i, over the Borel sets of T_i, and that the joint distribution is given by ρ. We assume that ρ is absolutely continuous with respect to the product measure $\prod_{i \in N} \rho_i$. Note that this implies that certain well-behaved types *of correlation* between the distribution of types for different voters are allowed in the model.

We assume three things about the distribution of preferences. First, we assume that u is *uniformly bounded* with respect to N, i.e., there exists a $D \in \Re$ such that for all $x \in X$ and $\tau \in \mathcal{T}$, $|u(x,\tau) - u(y,\tau)| < D$. This is essentially a restriction that rules out sequences of voters possessing arbitrarily strong policy preferences. In addition, uniform boundedness would follow from continuity of u and compactness of X and $\mathcal{T}$. Second, we assume that, for any set of voters N, there exists a unique policy, $x^* \in X$, which maximizes the expected sum of voters' utilities. Third, we assume the existence of a number M satisfying the following for all N:

$$M > \sup_{x \in X} \frac{-E_\tau\left[\sum_{i \in N}\left(u_i(x;\tau) - u_i(x^*;\tau)^2\right)\right]}{E_\tau\left[\sum_{i \in N}\left[u_i(x;\tau) - u_i(x^*;\tau)\right]\right]}. \tag{3.1}$$

This assumption rules out preferences which become arbitrarily "diverse" in relation to the optimality of any particular policy. Unfortunately, this restriction rules out some plausible preference profiles. On the other hand, we only need this condition for the final results of the Chapter, Theorem 46 and Corollary 47.

For notational simplicity, we drop the argument of $\eta_i(\tau_i)$, and just write η_i when there is no confusion. Also, η_{ij} is used to represent the j^{th} component of $\eta_i(\tau_i)$. All of the η_{ij} for $i \in N$, $j \in K_0$, and $\tau_i \in \mathcal{T}$ are assumed to be independently distributed absolutely continuous random variables with full support, each with a cumulative density function that is twice continuously differentiable. We assume that the $\eta_{ij}(\tau_i)$ are identically distributed for all $i \in N$, $j \in K$, and $\tau_i \in \mathcal{T}$. However, we allow for η_{i0} to have a different distribution than η_{ij} to allow for costs or benefits of voting. Any joint distribution ρ on T satisfying all of the above conditions is said to be *admissible*. Let μ be the common mean of η_{ij} for $j \in K$, μ_0 be the mean of η_{i0}, and $c = \mu - \mu_0$. Then c is the expected cost of voting.

We now define a game, in which the candidates each simultaneously choose policy positions in X and then, after observing the candidate policy positions, the voters vote for a candidate. Thus, the strategy set Y_i for candidate $i \in K$ is $Y_i = X$, and the set of strategy profiles for the candidates is $Y = \Pi_{i \in K} Y_i$. The strategy set S_i for voter $i \in N$ is the set of functions $s_i : Y \times T_i \to K_0$, and the set of strategy profiles for the voters is $S = \Pi_{i \in N} S_i$. We use the notation $S_{-i} = \Pi_{j \neq i} S_j$, and $s_{-i} \in S_{-i}$ to represent strategy profiles for all voters except voter i, with similar notation for candidates.

Given a strategy choice $y = (y_1, \ldots, y_k) \in Y$ of the candidates, and $s = (s_1, \ldots, s_n) \in S$ of the voters, define for any $j \in K_0$, and $t \in T^n$

$$V_j(y, s; t) = \frac{1}{n} \left| \{ i \in N : s_i(y, t_i) = j \} \right| \tag{3.2}$$

to be the proportion of the vote for j, and

$$W(y, s; t) = \{ j \in K : j \in \arg\max_{l \in K} (V_l(y, s; t)) \} \tag{3.3}$$

to be the set of winners of the election. For any $J \subseteq K$, write

$$P_J(y, s; t_i) = \Pr[\{ t_{-i} \in T_{-i} : W(y, s; t) = J \}]. \tag{3.4}$$

to be the probability of a first place tie among the candidates J. We assume that a fair lottery is used to select a winner when there is a tie, so that we can define voter utilities over subsets $J \subseteq K$ by

$$v_J(y, \tau_i) = \frac{1}{|J|} \sum_{j \in J} u(y_j, \tau_i). \tag{3.5}$$

The payoff to voter $i \in N$ of type $t_i = (\tau_i, \eta_i)$ from the strategy $(y, s) \in Y \times S$ is defined to be:

$$U(y, s, t_i) = \sum_{J \subseteq K} P_J(y, s; t_i) \cdot v_J(y, \tau_i) + \eta_{i s_i(y, t_i)} \tag{3.6}$$

In other words, a voter voting for candidate $j = s_i(y, t_i)$ receives the expected utility of the policy of the winning candidate, plus a payoff disturbance η_{ij} that is associated with the vote, $j \in K_0$ that the voter makes. We write $U(j; y, s, t_i) = U(y, (j, s_{-i}); t_i)$

for the utility that voter i of type t_i gets from voting for strategy j, given y, and $s_{-i} \in S_{-i}$. Since $P_J(y, s; t_i)$ is a function of t_i only through s_i, it follows that $P_J(y, (j, s_{-i}); t_i)$ is independent of t_i . So we write $P_J(y, (j, s_{-i})) = P_J(y, (j, s_{-i}); t_i)$. Then, we can write for all $j \in K_0$,

$$U(j; y, s, t_i) \quad = \quad \bar{U}(j; y, s, \tau_i) + \eta_{ij} \tag{3.7}$$

where

$$\bar{U}(j; y, s, \tau_i) = \sum_{J \subseteq K} P_J(y, (j, s_{-i})) \cdot v_J(y, \tau_i) \tag{3.8}$$

is the expected utility to voter i of type τ_i of voting for candidate j, unconditioned on the payoff disturbance, η_{ij}.

It follows from McKelvey and Ordeshook (1972) that the difference in the expected utility of voting for j over abstaining can be written in the form:[2]

$$\bar{U}(j; y, s, \tau_i) - \bar{U}(0; y, s, \tau_i) = \sum_{k \neq j} \delta_i^{jk}(y, s) \cdot [u(y_j, \tau_i) - u(y_k, \tau_i)] \tag{3.9}$$

where $\delta_i^{jk}(y, s)$ is the *pivot probability* for j over k:

$$\delta_i^{jk}(y, s) = \sum_{j,k \in J \subseteq K} \frac{1}{|J|} \left(q_J^0 + \frac{q_J^j}{|J| - 1} \right) \tag{3.10}$$

where we use the shorthand $q_J^k = P_J(y, (k, s_{-i}))$. The pivot probability is the probability that by voting for j rather than abstaining, voter i changes the outcome from a win for k to a win for j . It then follows from Equation 3.9 that the difference in

[2]Equation (3.9) follows by reversing the order of summation in the expression for $(E^j - E^0)$ of the Theorem on p. 49 of (1972).

expected utility of voting for j over l is:

$$\bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) = \left(\delta_i^{jl}(y, s) + \delta_i^{lj}(y, s)\right) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)] \quad (3.11)$$

$$+ \sum_{k \neq j,l} \left\{ \begin{array}{c} \delta_i^{jk}(y, s) \cdot [u(y_j, \tau_i) - u(y_k, \tau_i)] \\ +\delta_i^{lk}(y, s) \cdot [u(y_k, \tau_i) - u(y_l, \tau_i)] \end{array} \right\} \quad (3.12)$$

which, for the case of two candidates, $K = \{j, l\}$, reduces to

$$\bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) = \left(\delta_i^{jl}(y, s) + \delta_i^{lj}(y, s)\right) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)] \quad (3.13)$$

To define the candidate payoff functions, we first define $V_j(y, s)$ to be the expected proportion of the votes for candidate j at the profile (y, s):

$$V_j(y, s) = E_t [V_j(y, s; t)] = \frac{1}{n} E_t [|\{i \in N : s_i(y, t_i) = j\}|] \quad (3.14)$$

Then we define the payoff to candidate j to be the *margin of expected victory* $\widehat{V}_j$, defined by:

$$\widehat{V}_j(y, s) = V_j(y, s) - \max_{l \in N - \{j\}} V_l(y, s) \quad (3.15)$$

Remark 39 Any voter with unbounded utility would be subject to the St. Petersburg paradox: If x_k is chosen to satisfy $u(x_k, \tau_i) > 2^k$, for $k = 1, 2, \ldots$, the voter would not trade the lottery that gives prize x_k with probability $\frac{1}{2^k}$ for any x. Similarly, if the x_k satisfy $u(x_k, \tau_i) < -2^k$, they would not accept the lottery for any x. Thus, bounded utility for any one voter is implied if the voter is not subject to the St. Petersburg paradox. The uniform boundedness condition requires further that there be a common maximum and minimum bound across all voters.

Remark 40 Note that our assumptions do not preclude atoms in the marginal distribution of ρ over $\mathcal{T}$. The requirement that ρ be atomless is automatically satisfied via the assumptions that are imposed by admissibility on the distribution of the η_i's. Thus, our assumption of admissibility of ρ encompasses on the one hand the classical *framework*, in which all voter ideal points are known and common knowledge, and on the other hand, models such as that of Ledyard, in which all voter types are independent and drawn *i.i.d* from a common distribution on voter types. The classical framework arises if we let the marginal distribution of ρ on $\mathcal{T}$ be discrete.

Remark 41 The assumption that the distribution of the η_{ij} are *i.i.d* with respect to voters is an implicit normalization of utility functions. This is important in interpreting the main theorem, since the weights that individuals are given in the social utility function is determined by this normalization.

3.3 Voter Equilibrium

In this section, we consider the voter equilibrium to the game defined by equation 3.7 conditional on fixed candidate positions, $y \in Y$. Since the candidate positions are fixed, the strategy space for the voter reduces from S_i (the set of functions s_i : $Y \times T_i \to K_0$) to the set of functions of the form $s_i(y, \cdot) : T_i \to K_0$. We write $S_i(y)$ to designate this conditional strategy space, and $S(y)$ to designate the set of profiles of conditional strategies.

For any fixed $y \in Y$, we define a *voter equilibrium for* y to be a pure strategy Bayesian Nash equilibrium (BNE) to the voter game defined by (3.7) over the strategy space $S(y)$. This is any profile, $s \in S(y)$, in which voters always choose an action that maximizes expected utility conditional on their type. Thus, s is a voter equilibrium for y if for all $i \in N$, $t_i \in T_i$, and $j \in K_0$,

$$\begin{aligned} s_i(y, t_i) = j \quad &\Leftrightarrow \quad U(j; y, s, t_i) = \max_{l \in K_0} U(l; y, s, t_i) \\ &\Leftrightarrow \quad \bar{U}(j; y, s, \tau_i) + \eta_{ij} = \max_{l \in K_0} \left[\bar{U}(l; y, s, \tau_i) + \eta_{il}\right] \end{aligned} \tag{3.16}$$

Note the structure of the payoffs is exactly the same as used in McKelvey and Palfrey (1998) in defining the agent quantal response equilibrium (AQRE) for extensive form games. So as long as the distribution of the errors, η_{ij} is admissible, a Bayes Nash equilibrium to the voter game is exactly the same as an AQRE to the game.

Note further that in any voter equilibrium for y, except on a set of measure zero, the strategy $s_i(y, t_i)$ depends on i only through t_i. So we can drop the subscript on s without loss of generality.

Proposition 42 *For any $y \in Y$, there exists a voter equilibrium for y.*

Proof: This is a game of incomplete information, with action spaces $A_i = K_0$ and type space T_i for each $i \in N$. The action spaces are finite, and the distribution of types is independent across individuals. Thus, we can apply Theorem 1 of Milgrom and Weber (1985) to conclude that there exists an equilibrium in distributional strategies. Further, since the distribution of player i's types, ρ_i, is assumed atomless, it follows from Theorem 4 in the same paper that the equilibrium can be purified to be in pure strategies. ∎

Of particular interest is the average behavior of a voter i of type t_i, after integrating out η_i. For any $s_i(y, \cdot) \in S_i(y)$, define $\bar{s}_i(y, \cdot) : \mathcal{T} \to \Delta^{K_0}$, as the marginal distribution of s_i with respect to η_i: for any $\tau_i \in \mathcal{T}$ and $j \in K_0$,

$$\bar{s}_i(y, \tau_i)(j) = \Pr[\eta_i : s_i(y, (\tau_i, \eta_i)) = j]. \tag{3.17}$$

We have assumed that the η_{ij} are independently distributed, for all i, j and τ_i, and identically distributed for all $j \in K$. Let $H(\cdot)$ be the cumulative distribution function of η_i, i. e., $H(w) = Pr[\eta_{ij} \leq w_j \text{ for all } j \in K_0]$ for $w \in \Re^{K_0}$. And let $G_j(\cdot)$ be the cumulative distribution function of $\zeta \in \Re^K$, where $\zeta_l = \eta_{il} - \eta_{ij}$ for $l \in K - \{j\}$, and $z_j = \eta_{i0} - \eta_{ij}$. Thus,

$$G_j(z) = \Pr[\eta_{i0} - \eta_{ij} \leq z_j \text{ and } \eta_{il} - \eta_{ij} \leq z_l \text{ for all } l \neq j] \tag{3.18}$$

for any $z \in \Re^K$. Under the assumptions we have made on the η_{ij}, for all $j \in K$, both $H(w)$ and $G_j(z)$ are twice continuously differentiable and strictly increasing in all arguments, and everywhere positive. Thus, if s is a Bayes Nash equilibrium, applying *equation (3.16)*, for $j \in K$,

$$\begin{aligned}\bar{s}_i(y, \tau_i)(j) &= \Pr[\bar{U}(j; y, s, \tau_i) + \eta_{ij} = \max_{l \in K_0} [\bar{U}(l; y, s, \tau_i) + \eta_{il}]] \\ &= \Pr[\eta_{il} - \eta_{ij} \leq \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) \text{ for all } l \in K_0 - \{j\}] \\ &= G_j(\bar{\mathbf{U}}^j(y, s, \tau_i)). \end{aligned} \tag{3.19}$$

where $\bar{\mathbf{U}}^j(y, s, \tau_i)$ is a vector in $\Re^K$ with components $\bar{\mathbf{U}}^j_l(y, s, \tau_i) = \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i)$ for $l \neq j$, and $\bar{\mathbf{U}}^j_j(y, s, \tau_i) = \bar{U}(j; y, s, \tau_i) - \bar{U}(0; y, s, \tau_i)$.

Example: One example of the above is the logit AQRE, where the density functions of $w_0 = \eta_{i0} + c$ and $w_j = \eta_{ij}$ for $j \in K$ follow a type one extreme value distribution, $H_j(w_j) = \exp[-\exp[-\lambda w_j]]$. Thus, with independence, we have $H(w) = \prod_j H_j(w_j)$. This leads to the logistic formula $G_j(z) = \frac{1}{1+\exp\lambda(c+z_j)+\sum_{l\neq j}\exp(\lambda z_l)}$. In this case, for fixed λ, we get:

$$\begin{aligned}\bar{s}_i(y, \tau_i)(j) &= G_j(\bar{\mathbf{U}}^j(y, s, \tau_i)) \\ &= \frac{1}{\begin{array}{c}1 + \exp\left[\lambda \cdot \left(c + \bar{U}(0; y, s, \tau_i) - \bar{U}_i(j; y, s, \tau_i)\right)\right] \\ + \sum_{l \neq j} \left(\exp\left[\lambda \cdot \left(\bar{U}(l; y, s, \tau_i) - \bar{U}(j; y, s, \tau_i)\right)\right]\right)\end{array}},\end{aligned}$$

and in the case of two candidates, where $K = \{j, l\}$,

$$\bar{s}_i(y, \tau_i)(j) = \frac{1}{\begin{array}{c}1 + \exp\left(\lambda \cdot \left(c + \delta^{jl}(y, s) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)]\right)\right) \\ + \exp\left(\lambda \cdot (\delta^{jl}(y, s) + \delta^{lj}(y, s)) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)]\right)\end{array}}.$$

We now show that for fixed candidate positions at $y \in Y$, and for any voter equilibrium, that all pivot probabilities go to zero and the probability of voting for any two candidates in K becomes equal as $n \to \infty$. The reason for this result is simple: one's vote only matters when it is pivotal.[3] Thus, one's vote only matters

[3] The logic of pivotal voting is explained in the voting literature. See e.g., Myerson and Weber (1995).

when the other voters are either evenly split between the two top candidates or when the vote difference between the two top candidates differs by one vote. As n grows large, this becomes a very low probability event. Thus, in general, one's vote doesn't make a difference very often. This implies that voters effectively become indifferent with respect to which candidate they vote for as $n \to \infty$. We formalize the above in the following proposition:

Proposition 43 *Assume u is uniformly bounded. Fix $y \in Y$, and for each integer n, let ρ^n be any admissible joint distribution over $\Pi_{i=1}^n T_i$, and let s^n be any AQRE for the voters. Then for any $j, l \in K$ and $i, k > 0$,*

(a) $\lim_{n\to\infty} \delta_i^{jl}(y, s^n) = 0$ *and*

(b) $\lim_{n\to\infty} \delta_i^{jl}(y, s^n)/\delta_k^{jl}(y, s^n) = 1$

(c) $\lim_{n\to\infty} \delta_i^{jl}(y, s^n)/\delta_i^{lj}(y, s^n) = 1$

(d) $\lim_{n\to\infty}[\bar{s}_i^n(y, \tau_i)(j) - \bar{s}_i^n(y, \tau_i)(l)] = 0.$

Further, in all cases, the convergence is uniform. I. e., for any $\varepsilon > 0$, there is an n_ε such that for all $i, k, j, l, y, \rho^n, s^n$ if $n > n_\varepsilon$, $\delta_i^{jl}(y, s^n) < \varepsilon$, $\left|\delta_i^{jl}(y, s^n)/\delta_k^{jl}(y, s^n) - 1\right| < \varepsilon$, and $\bar{s}_i^n(y, \tau_i)(j) - \bar{s}_i^n(y, \tau_i)(l) < \varepsilon$.

To prove the proposition, we need a Lemma.

Lemma 44 *Fix $\varepsilon^* > 0$, and let $\mathcal{Z}^n$ be the set of sequences $Z = (Z_1, \ldots, Z_n)$ of independent random vectors $Z_i \in \Re^K$ of the form*

$$Z_i = \alpha_j \ w.\ p.\ p_{ij}$$

where α_j is the j^{th} unit basis vector in $\Re^K$, and $p \in \left(\Delta^{K_0}\right)^n$ satisfies $p_{ij} \geq \varepsilon^*$ for all i, j. For any $J \subseteq K$, define

$$B_J = \{z \in \Delta^K : z_j = z_k > z_l \text{ for all } j, k \in J,\ l \notin J\}.$$

Write $\bar{Z} = \sum_i Z_i$, *and define*

$$\delta_J^{n*} = \max_{Z \in \mathcal{Z}^n} \Pr[\bar{Z} \in B_J] \tag{3.20}$$

Then for any $J \subseteq K$ *with* $|J| \geq 2$

(a) $\lim_{n\to\infty} \delta_J^{n*} = 0$

(b) $\lim_{n\to\infty} \delta_{J'}^{n*}/\delta_J^{n*} = 0$ *for any* $J \subsetneq J'$

Proof: An element $Z = (Z_1, \dots, Z_n) \in \mathcal{Z}^n$ consists of independent, but not identically distributed random vectors, and is characterized by a vector $p = (p_1, \dots, p_n)$, where $p_i = (p_{i0}, p_{i1}, \dots, p_{iK}) \in \Delta^{K_0}$. The mean of Z_i is $\mu_i = (p_{i1}, \dots, p_{iK})$ which consists of all but the first component of p. Pick $Z^n = (Z_1^n, \dots, Z_n^n) \in \mathcal{Z}^n$ to attain the maximum in Equation 3.20. Since $\Pr[\bar{Z} \in B_J]$ is continuous as a function of p, which ranges over a compact set, it follows that such a δ_J^{n*} and Z^n exist. Define V_{ni} to be the variance covariance matrix of Z_i^n, and $X_i^n = Z_i^n - \mu_i$. Set $V_n = \frac{1}{n}\sum_i V_{ni}$ and $T_n^2 = V_n^{-1}$. From our assumption that $p_{ij} > \varepsilon^*$ for all $j \in K_0$, it follows that V_n is strictly positive definite and hence invertible. Then

$$\begin{aligned}
\delta^{n*} &= \Pr[\bar{Z}^n \in B_J] \\
&= \Pr\begin{bmatrix} \sum_i Z_{ij}^n - \sum_i Z_{ik}^n = 0 \text{ for } j,k \in J, \text{ and} \\ \sum_i Z_{ij}^n - \sum_i Z_{il}^n > 0 \text{ for } j \in J, l \notin J \end{bmatrix} \\
&= \Pr\begin{bmatrix} \sum_i \left(X_{ij}^n - X_{ik}^n\right) = \sum_i (p_{ik} - p_{ij}) \text{ for } j,k \in J, \text{ and} \\ \sum_i \left(X_{ij}^n - X_{il}^n\right) > \sum_i (p_{il} - p_{ij}) \text{ for } j \in J, l \notin J \end{bmatrix} \\
&= \Pr\begin{bmatrix} \frac{T_n}{\sqrt{n}}\sum_i \left(X_{ij}^n - X_{ik}^n\right) = \frac{T_n}{\sqrt{n}}\sum_i (p_{ik} - p_{ij}) \text{ for } j,k \in J, \text{ and} \\ \frac{T_n}{\sqrt{n}}\sum_i \left(X_{ij}^n - X_{il}^n\right) > \frac{T_n}{\sqrt{n}}\sum_i (p_{il} - p_{ij}) \text{ for } j \in J, l \notin J \end{bmatrix}
\end{aligned} \tag{3.21}$$

But now the X_i^n form a triangular array where each random variable X_i^n has zero mean, and for each n, the X_i^n are independent. Further, writing Q_i^n for the cumulative density function of X_i^n, the random vectors satisfy the following multivariate

Lindeberg condition: For every $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_i \int_{\|T_n X_i\| > \epsilon\sqrt{n}} \|T_n X_i\|^2 \, dQ_i^n(X) = 0 \tag{3.22}$$

To see this, note that Z_i^n is in the simplex Δ^K. Hence, $\|X_i^n\| \leq 2$. The probability that $Z_{ij}^n = 1$ is $p_{ij} \geq \varepsilon^*$. Further, the variances and covariance of V_{ni} are all uniformly bounded away from zero and one, since $p_{ij} \geq \varepsilon^*$ for all i, j. Thus, the same will be true of V_n. So V_n will be invertible, and for any ϵ, we can pick large enough n so that $\|T_n X_i\| < \epsilon\sqrt{n}$. So each term in the summation of Equation 3.22 goes to zero with n, which establishes (3.22). It follows by the multivariate version of the central limit theorem for triangular arrays (see Bhattacharya and Rao, (1986), Corollary 18.2, p 183) that the distribution of $\frac{1}{\sqrt{n}} T_n \sum_i X_i^n$ converges weakly to a multivariate unit normal distribution. Hence the probability it falls in a subset of any lower dimensional subspace goes to zero. Thus, when $|J| \geq 2$, the right-hand side of Equation 3.21 converges to 0 with n. I. e., $\lim_{n\to\infty} \delta^{n*} = 0$, proving (a). To prove (b), we note that $B_{J'}$ describes a lower dimensional subspace than B_J. Hence, an argument similar to above shows that for all sequences, the $\Pr[\bar{Z} \in B_{J'}]$ goes to zero faster than $\Pr[\bar{Z} \in B_J]$ establishing the result. ∎

We now proceed to a proof of the proposition.

Proof: To prove (a), define $D = 2 \cdot (|K| - 1) \cdot \sup_{x,y,\tau} [u(x, \tau) - u(y, \tau)]$, and ε^* $= \min_{j\in K} G_j(-\mathbf{1} \cdot D)$, where $\mathbf{1} = (1, \dots, 1)$ is the unit vector of length $|K|$. By the assumptions we have made on the η_{ij}, $\varepsilon^* > 0$. Then from Equation 3.9, using the fact that $\delta_i^{jl} \leq 1$ for all i, j, k,we have $-D \leq \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) \leq D$ for all $j, l \in K$, which implies that $\bar{s}_i(y, \tau_i)(j) = G_j(\bar{\mathbf{U}}^j(y, s, \tau_i)) \geq G_j(-\mathbf{1} \cdot D) \geq \varepsilon^*$.

Now, given any sequence $\tau = (\tau_1, \dots, \tau_n)$ with $\tau_i \in \mathcal{T}$ for all $i > 0$, define the random variable

$$Z_{ni}(\tau_i) = \alpha_j \text{ if } s_i^n(y, (\tau_i, \eta_i)) = j$$

So $Z_{ni}(\tau_i) \in \mathcal{Z}^n$, with $p_{ij} = \bar{s}_\alpha^n(y, \tau_\alpha)(j)$.

Then, letting $(0, s^n_{-i})$ be the profile where the voter i abstains, and (j, s^n_{-i}) be the profile where voter i votes for candidate j, we have, from Equation 3.10:

$$\delta_i^{jl}(y, s^n) = \sum_{j,k \in J \subseteq K} \frac{1}{|J|} \left(q_J^0 + \frac{q_J^j}{|J| - 1} \right) \tag{3.23}$$

But, from Equation 3.4, for any $J \subseteq K$,

$$\begin{aligned} q_J^0 &= P_J(y, (0, s_{-i})) = \Pr[\{t_{-i} \in T_{-i} : W(y, s; t) = J\}] & (3.24) \\ &= E_{t_{-i}} \left[W(y, (0, s^n_{-i}); t_{-i}) = J \right] = E_{t_{-i}} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] & (3.25) \\ &= E_{\tau_{-i}} E_{\eta_{-i}} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] \leq E_{\tau_{-i}} [\delta_J^{n*}] = \delta_J^{n*}, & (3.26) \end{aligned}$$

where the inequality follows from the definition of δ_J^{n*} in Lemma 44. A similar argument shows the second term in Equation 3.23 is less than or equal to δ_J^{n*}. Thus, $\delta_i^{jl}(y, s^n) \leq \sum_{j,k \in J \subseteq K} \left(\frac{1}{|J|-1} \right) \delta_J^{n*} \leq \left(\sum_{j,k \in J \subseteq K} \frac{1}{|J|-1} \right) \delta^{n*}$, where $\delta^{n*} = \max_{J \subseteq K} \delta_J^{n*}$ By Lemma 44, $\lim_{n \to \infty} \delta^{n*} = 0$, which proves *(a)*. Since δ^{n*} is independent of i, j, l, y, the convergence is uniform in all arguments.

To show (b), for each $J \subseteq K$, we can write $P_J(y, (0, s_{-i})) = E_{t_{-i}} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right]$ the corresponding expression for voter j is $P_J(y, (0, s_{-j})) = E_{t_{-j}} \left[\sum_{l \neq j} Z_{nl}(\tau_l) \in B_J \right]$. But the RHS of these two expressions differ only by the i and j^{th} terms, and hence, by Lemma 44, both converge weakly to the same multivariate normal distribution. Hence, in the limit, the ratio of the two must approach one. The same argument applies to all terms in the sum in (3.23). Thus, the result follows. A similar argument suffices to establish (c).

To show (d), we have from Equation 3.17 that

$$\bar{s}_i^n(y, \tau_i)(j) = \Pr[\max_{l \neq j} \bar{U}(l; y, s^n, \tau_i) + \eta_{il} \leq \bar{U}(j; y, s^n, \tau_i) + \eta_{ij}].$$

Now, in the first part of the proposition we showed all pivot probabilities go to zero uniformly as n gets large. Hence, using Equation 3.12 we get that as $n \to \infty$, for

$j, l \in K$, $\bar{U}(l; y, s^n, \tau_i) - \bar{U}(j; y, s^n, \tau_i) \to 0$ uniformly in i, j, l, y, τ. But then we get

$$\begin{aligned} \lim_{n\to\infty} [\bar{s}_i^n(y,\, \tau_i)(j) - \bar{s}_i^n(y,\, \tau_i)(l)] &= \Pr[\max_{\alpha\neq j} \eta_{ia} - \eta_{ij} \leq 0] - \Pr[\max_{a\neq l} \eta_{ia} - \eta_{il} \leq 0] \\ &= G_j(0) - G_l(0) = 0. \end{aligned} \tag{3.27}$$

Since the convergence of $\bar{U}(l; y, s^n, \tau_i) - \bar{U}(j; y, s^n, \tau_i)$ is uniform in all arguments, it follows that the convergence in Equation 3.27 is also. ∎

Based on Proposition 43 (b), it follows that for large n, we can ignore the voter subscript on δ, and write $\delta_i^{jl}(y, s^n) = \delta^{jl}(y, s^n) = \delta^{lj}(y, s^n)$. Further, from Lemma 44, it follows that in any voter equilibrium, all ties involving three or more candidates will be small in relation to the two candidate ties. Recall the notation $q_J^k = P_J(y, (k, s_{-i}))$. Then for $J \subsetneq J'$,

$$\lim_{n\to\infty} q_J^k / q_{J'}^k = \lim_{n\to\infty} P_J(y, (k, s_{-i}))/P_{J'}(y, (k, s_{-i})) = 0$$

Hence, for large electorates, formula (3.10) for the pivot probability has the following approximation:

$$\delta_i^{jk}(y, s) = \sum_{j,k\in J\subseteq K} \frac{1}{|J|}\left(q_J^0 + \frac{q_J^j}{|J|-1}\right) \cong \frac{1}{2}\left(q_{\{j,k\}}^0 + q_{\{j,k\}}^j\right)$$

Remark 45 Note that the requirement that voters adopt a Bayesian equilibrium means that voters vote strategically in multi-candidate elections, Thus, a voter may rank $u(y_j, \tau) > u(y_l, \tau)$, and yet (even if the realization of the payoff disturbances is zero) vote for their second ranked alternative l over their first ranked alternative j if the pivot probability for the first ranked alternative is sufficiently low in relation to that for the second ranked alternative so that we have $\bar{U}(l; y, s^n, \tau_i) - \bar{U}(j; y, s^n, \tau_i) > 0$.

3.4 Candidate Equilibrium

This section examines the incentives of candidates competing for votes in a world populated by voters who play quantal response equilibrium strategies. We establish that for a large enough electorate, N, all candidates adopting the social optimum constitutes a global equilibrium. In the case of two candidates, the global equilibrium is unique. Our results hold regardless of the how the measure ρ changes as the size of the electorate increases, as long as the admissibility condition is met. More specifically, recall that admissibility required that the η_{ij} are *i.i.d.* with full support. We also assume that the distribution of the η_{ij} is independent of the size n of N.

For a fixed electorate, N, and measure ρ on $T = \prod_{i \in N} T_i$, let s be any strategy profile for the voters[4] such that for any candidate positions, $y \in Y$, $s(y, \tau)$ is a quantal response equilibrium for the voters, as described in the previous section. We use the notation

$$V_j(y) = V_j(y, s(y, \cdot)) = E_t\left[V_j(y, s(y, t); t)\right] \tag{3.28}$$

to represent the expected vote for the candidates j, assuming that the voters follow the strategy s in response. Then,

$$\begin{aligned} V_j(y) &= \frac{1}{n} E_t\left[|\{i \in N : s_i(y, t_i) = j\}|\right] \\ &= \frac{1}{n} E_\tau\left[E_\eta\left[|\{i \in N : s_i(y, \tau_i, \eta_i) = j\}|\right]\right] \\ &= \frac{1}{n} E_\tau\left[\sum_{i \in N} \bar{s}_i(y, \tau_i)(j)\right] = \frac{1}{n} \sum_{i \in N} E_{\tau_i}\left[\bar{s}_i(y, \tau_i)(j)\right]. \end{aligned} \tag{3.29}$$

We assume that candidates seek to maximize the margin of expected victory. So the payoff of candidate $j \in K$ at the profile (y, s) is given by:

$$\widehat{V}_j(y) = \widehat{V}_j(y, s) = V_j(y) - \max_{l \in N - \{j\}} V_l(y). \tag{3.30}$$

[4] To be technically correct, since we are considering N and ρ to be variables, we should subscript voter and candidate strategies on these variables. To simplify notation, we leave off these parameters.

Let

$$x_\rho^* = \arg\max_{x\in X} \sum_{i\in N} E_{\tau_i}[u(x,\tau_i)] \tag{3.31}$$

denote the *expected social optimum*. We assume for each N and ρ that such a point exists and is unique.

Theorem 46 *Let u be uniformly bounded. There exists an integer n^* such that for any set of voters N with $|N| = n > n^*$, and any admissible ρ on $T = \prod_{i\in N} T_i$, $y^* = (x_\rho^*, \ldots, x_\rho^*)$ constitutes a global equilibrium under the margin of expected victory: for any $j \in K$ and $y_j \in X$, $\hat{V}_j(y) = \hat{V}_j(y_j, y_{-j}^*) \leq \hat{V}_j(y^*)$, with the weak inequality becoming strict whenever $y_j \neq x^*$.*

Proof: For any set of voters N, and admissible ρ, let $y = (y_j, y_{-j}^*)$, where $y_l^* = x_\rho^*$ for all $l \neq j$ and $y_j \neq x_\rho^*$. We first show that for large enough n, $V_j(y) = V_j(y_j, y_{-j}^*) \leq V_j(y^*)$.

For $z \in \Re^K$, write $Q(z) = G_j(z)$, where G_j is as defined in Equation 3.18. Given an individual $i \in N$, and using equations (3.9) and (3.12), the probability of a vote for candidate j is given by

$$\begin{aligned} s_i(y,\tau_i)(j) &= \Pr\left[\max_{l\in K_0-\{j\}} [U(l;y,s,t_i) - U(j;y,s,t_i)] \leq 0\right] \\ &= \Pr\left[\begin{array}{c} \eta_{ik} - \eta_{ij} \leq \Delta_i^k(y,s)\cdot[u(y_j,\tau_i) - u(x_\rho^*,\tau_i)] \text{ for } k \in K - \{j\} \\ \text{and } \eta_{i0} - \eta_{ij} \leq \Delta_i^j(y,s)\cdot[u(y_j,\tau_i) - u(x_\rho^*,\tau_i)] \end{array}\right] \\ &= Q(\Delta_i(y,s)\cdot[u(y_j,\tau_i) - u(x_\rho^*,\tau_i)]) \end{aligned} \tag{3.32}$$

where $\Delta_i(y,s) = (\Delta_i^1(y,s), \ldots, \Delta_i^k(y,s))$, $\Delta_i^l(y,s) = 2\delta_i^{lj}(y,s) + \sum_{\alpha\neq j,l} \delta_i^{\alpha j}(y,s)$, for all $l \in K - \{j\}$, and $\Delta_i^j(y,s) = \sum_{\alpha\neq j} \delta_i^{j\alpha}(y,s)$.

Using Equation 3.29 we can express the vote for candidate j as

$$V_j(y) = \frac{1}{n} \sum_{i\in N} E_{\tau_i}[\bar{s}_i(y,\tau_i)(j)] \tag{3.33}$$

Then, from Equation 3.32, we have that

$$V_j(y) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(\Delta_i(y,s)\cdot[u(y_j,\tau_i) - u(x^*_\rho,\tau_i)])\right]$$

Without loss of generality, we can assume utility functions are normalized with $u(x^*,\tau_i) = 0$ for all $i \in N$ and $\tau_i \in T$. Write $u_i = u(y_j,\tau_i) \in \mathbb{R}$, and $\Delta_i = \Delta_i(y,s)$. Then, the above can be written as:

$$V_j(y) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(\Delta_i(y,s)\cdot u(y_j,\tau_i))\right] = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(\Delta_i u_i)\right] \tag{3.34}$$

Using parts (b) and (c) of Proposition 43, normalize the Δ_i by Δ_1 in the following manner. For $i \in N$, let

$$\chi_i = \left(\frac{\Delta_i^1}{\Delta_1^1},\ldots,\frac{\Delta_i^k}{\Delta_1^k}\right)^T,$$

and

$$D = \begin{bmatrix} \Delta_1^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Delta_1^k \end{bmatrix}.$$

It is easily shown that $\Delta_i^j > 0$ for all $i \in N$ and $j \in K$, so that χ_i is well defined. Then, applying Taylor's theorem, we can write

$$\begin{aligned} V_j(y) - V_j(y^*) &= \frac{1}{n}\sum_{i\in N}\left\{E_{\tau_i}\left[Q\left(D\cdot\chi_i\cdot u_i\right)\right] - E_{\tau_i}\left[Q\left(0\right)\right]\right\} \\ &= \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[(D\cdot\chi_i\cdot u_i)^T Q'\left(\mathbf{0}\right) + \frac{1}{2}(\Delta_i\cdot u_i)^T Q''\left(z_i(y)\right)\left(\Delta_i\cdot u_i\right)\right] \end{aligned} \tag{3.35}$$

where $z_i(y) = \alpha\mathbf{0} + (1-\alpha)(0,\ldots,u_i,0,\ldots,0)$ for some $\alpha \in (0,1)$ for each $i \in N$, and $\mathbf{0}$ denotes a k dimensional vector of zeros.

Now, by Proposition 43, it follows that for any $\epsilon > 0$, we can find a value $n*$ such that $\max D < \epsilon$, $\max_{i\in N}[\max[(\chi_i - \mathbf{1})]] < (1-\epsilon)$, and $\max_{i\in N}[\max[(\Delta_i]] < \epsilon$ for all $n > n*$. Using these facts and continuing the derivation of $V_j(y) - V_j(y^*)$,

$$\begin{aligned}
V_j(y) - V_j(y^*) &= \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q'(\mathbf{0})\,D\cdot\chi_i\cdot u_i + \frac{1}{2}(\Delta_i\cdot u_i)^T Q''(z_i(y))\,(\Delta_i\cdot u_i)\right] \\
&= \frac{1}{n}Q'(\mathbf{0})\cdot D\cdot\sum_{i\in N} E_{\tau_i}\left[\mathbf{1}\cdot u_i + (\chi_i - \mathbf{1})\cdot u_i\right] \\
&\quad + \frac{1}{2n}\sum_{i\in N} E_{\tau_i}\left[\Delta_i^T\cdot Q''(z_i(y))\cdot\Delta_i\cdot u_i^2\right] \\
&\le \frac{k}{n}(1-\epsilon)\epsilon Q^*\sum_{i\in N} E_{\tau_i}[u_i] + \frac{k^2}{2n}\epsilon^2 Q^{**}\sum_{i\in N} E_{\tau_i}\left[u_i^2\right] \qquad (3.36)
\end{aligned}$$

where $Q'(\mathbf{0})$ is a k dimensional vector consisting of the gradient of Q evaluated at $\mathbf{0}$, $Q''(\mathbf{0})$ is a $k\times k$ symmetric matrix of second partial derivatives of Q evaluated at $\mathbf{0}$, Q^* is the smallest element of $Q'(\mathbf{0})$, Q^{**} is defined as

$$Q^{**} = \sup_{z\in\mathbb{R}^k}[Q''(z)],$$

and $\mathbf{1}$ represents a k dimensional vector of ones.

We now want to show that there exists n^* such that for all N with $n > n^*$, the right-hand side of Equation 3.36 becomes less than zero for all $y \neq x^* \in X$. For a given $y \neq x^*$,

$$\begin{aligned}
\frac{k}{n}(1-\epsilon)\epsilon Q^*\sum_{i\in N} E_{\tau_i}[u_i] + \frac{k^2}{2n}\epsilon^2 Q^{**}\sum_{i\in N} E_{\tau_i}\left[u_i^2\right] &< 0 \\
\frac{k^2}{2n}\epsilon^2 Q^{**}\sum_{i\in N} E_{\tau_i}\left[u_i^2\right] &< -\frac{k}{n}(1-\epsilon)\epsilon Q^*\sum_{i\in N} E_{\tau_i}[u_i] \\
\frac{-\sum_{i\in N} E_{\tau_i}\left[u_i^2\right]}{\sum_{i\in N} E_{\tau_i}[u_i]} &< \frac{2(1-\epsilon)}{k\epsilon}\frac{Q^*}{Q^{**}},
\end{aligned}$$

The inequality in Equation 3.37 is satisfied for sufficiently small $\epsilon > 0$. Of course, this is for a given $y \neq x^*$. In order to satisfy Equation 3.37 for all $y \neq x^*$, we must take the supremum of the left-hand side over all $y \neq x^*$. This supremum is defined to be finite and denoted by M in Equation 3.1, resulting in the following requirement for Equation 3.37 to be satisfied for all $y \neq x^*$:

$$M \;<\; \frac{2(1-\epsilon)}{k\epsilon}\frac{Q^*}{Q^{**}}. \tag{3.37}$$

As with Equation 3.37, for ϵ sufficiently small, Equation 3.37 is satisfied. Thus, for any $y_j \in Y_j$, $V_j(y) = V_j(y_j, y^*_{-j}) \leq V_j(y^*)$ with strict inequality whenever $y_j \neq x^*_\rho$.

Next, we show that for some $l \neq j$, $V_l(y_j, y^*_{-j}) \geq V_l(y^*)$. We pick $l \in K - \{j\}$ for which $\delta^{jl}(y, s)$ is maximized. For $z \in \Re^K$, write $Q(z) = G_l(z)$, where G_l is as defined in equation (3.18). Then we have

$$\begin{aligned}
s_i(y,\tau_i)(l) &= \Pr\left[\begin{array}{c} U\,(0; y, s, \tau_i) - U(l; y, s, \tau_i) \leq 0, \text{and} \\ U(j; y, s, \tau_i) - U\,(l; y, s, \tau_i) \leq 0, \text{and} \\ \max_{k\in K-\{l,j\}}\left[U(k; y, s, \tau_i) - U(l; y, s, \tau_i)\right] \leq 0, \end{array}\right] \\
&= \Pr\left[\begin{array}{l} \eta_{i0} - \eta_{il} \leq \Delta^l_i(y,s)\cdot[u(x^*_\rho, \tau_i) - u(y_\alpha, \tau_i)], \text{and} \\ \eta_{ij} - \eta_{il} \leq \Delta^j_i(y,s)\cdot[u(x^*_\rho, \tau_i) - u(y_j, \tau_i)], \text{and} \\ \eta_{ik} - \eta_{il} \leq \max_{k\in K-\{l,j\}}\left(\Delta^k_i(y,s)\cdot[u(x^*_\rho, \tau_i) - u(y_j, \tau_i))\right] \end{array}\right] \\
&= Q(\Delta_i(y,s)\cdot[u(x^*_\rho, \tau_i) - u(y_j, \tau_i)])
\end{aligned}$$

where $\Delta_i(y, s) = (\Delta^1_i(y, s), \ldots, \Delta^k_i(y, s))$, with $\Delta^l_i(y, s) = \sum_{\alpha\neq l} \delta^{l\alpha}_i(y, s)$, $\Delta^j_i(y, s) = 2\delta^{lj}_i(y, s) + \sum_{\alpha\neq j,l} \delta^{j\alpha}_i(y, s)$, and $\Delta^k_i(y, s) = \delta^{jl}_i(y, s) - \delta^{kj}_i(y, s)$ for all $k \in K - \{l, j\}$.

Using equation (3.29) we can express the vote for candidate l as

$$V_l(y) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[s_i(y,\tau_i)(l)\right] = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(\Delta_i(y,s)\cdot[u(x^*_\rho,\tau_i) - u(y_j,\tau_i)])\right] \tag{3.38}$$

As above, we can assume utility functions are normalized with $u(x^*, \tau_i) = 0$ for all $i \in N$ and $\tau_i \in T$. As before, write $u_i = u(y_j, \tau_i)$, and $\Delta_i = \Delta_i(y, s)$. Then, the above can be written as:

$$V_l(y) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(-\Delta_i(y,s)\cdot u(y_j,\tau_i))\right] = \frac{1}{n}\sum_{i\in N} E_{\tau_i}\left[Q(-\Delta_i u_i)\right] \tag{3.39}$$

Note that the above takes exactly the same form as equation (3.34) above, with the exception of the negative sign. Consequently, an analogous argument to that in (3.36) establishes that we can find large enough n so that $V_l(y) - V_l(y^*)$ is positive. Thus, for any $y_j \in Y_j$, $V_l(y) = V_l(y_j, y^*_{-j}) \geq V_l(y^*)$ with strict inequality whenever $y_j \neq x^*_\rho$. We have shown that $V_j(y_j, y^*_{-j}) \leq V_l(y^*)$ and $V_l(y_j, y^*_{-j}) \geq V_l(y^*)$. So $\hat{V}_j(y_j, y^*_{-j}) \leq \hat{V}_j(y^*)$. So y^* is a global equilibrium for the objective function $\hat{V}$. ∎

For the case of two candidates, the above theorem can be strengthened:

Corollary 47 *If $k = 2$, then the equilibrium found in Theorem 46 is unique.*

Proof: Suppose there is another equilibrium, y. Then for at least one candidate j, $y_j \neq x^*$. Assume W.L.O.G. that $j = 2$. By Theorem 46, $\hat{V}_1(y_1, y_2) \geq \hat{V}_1(x^*, y_2) > 0$. Hence, $\hat{V}_2(y_1, y_2) < 0$. But this can not be an equilibrium for candidate 2, since $\hat{V}_2(y_1, x^*) \geq 0 > \hat{V}_2(y_1, y_2)$. This yields a contradiction. Hence the equilibrium is unique. ∎

Note that in the equilibrium defined by Theorem 46, that $y^*_j = y^*_l = x^*_\rho$ for all $j, l \in K$. Hence, we have $u(y^*_j, \tau_i) = u(y^*_l, \tau_i)$ for all $j, l \in K$. Thus, the level of abstention in equilibrium is determined by $V_0(y^*, s) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}[s_i(y^*, \tau_i)(0)]$. But

$$\begin{aligned}
\bar{s}_i(y^*, \tau_i)(0) &= \Pr\left[\max_{l\in K}\left[\bar{U}(l; y^*, s, t_i) - \bar{U}(0; y^*, s, t_i) + \eta_{il} - \eta_{i0}\right] \leq 0\right] \\
&= \Pr\left[\max_{l\in K}\left[\sum_{\alpha\neq l}\delta^{l\alpha}(y, s)\cdot[u(y^*_l, \tau_i) - u(y^*_\alpha, \tau_i)] + \eta_{il} - \eta_{i0}\right] \leq 0\right] \\
&= \Pr\left[\eta_{i0} \geq \max_{l\in K}[\eta_{il}]\right].
\end{aligned}$$

For example, if $c_i = 0$ for all $i \in N$, then under the assumptions we have made, all of the η_{il} for $l \in K_0$ are *i. i. d.* Hence the above evaluates to $\frac{1}{K+1}$. It follows that

$$V_0(y^*) = \frac{1}{n}\sum_{i\in N} E_{\tau_i}[s_i(y^*, \tau_i)(0)] = \frac{1}{K+1}.$$

So that in a two candidate election, one would obtain equilibrium turnout of about two thirds of the electorate. Of course, the above calculation would be very sensitive to the assumed distribution of costs of voting.

Thus, asymptotically we find that the social optimum is a global equilibrium so long as preferences are uniformly bounded.

3.5 Conclusions

We have provided a general framework for probabilistic spatial voting models in large electorates. In particular, we have extended equilibrium results of Coughlin, Ledyard, and other researchers to spaces of arbitrary finite dimensionality and elections with both abstention and arbitrary numbers of candidates. In addition, our model allows for strategic behavior by the voters.

As an aside, our model is agnostic as to the cause of probabilistic choice. The probabilistic choice in a QRE model can be assumed to arise either as the result of rational behavior under payoff disturbances (as we have modeled it here), or as the result of boundedly rational behavior. A key point to note in interpreting our results is that, with respect to the distributions of the alternative-specific payoff disturbance, we have assumed only that these distributions possess full support and are independently and identically distributed, while allowing for the abstention-specific shock to possess a nonidentical, independent distribution. In particular, we do not require that these payoff shocks be in any sense large. Thus, our results demonstrate the existence of an asymptotic equilibrium at the social welfare optimum in a relatively large class of probabilistic voting models.

Chapter 4 Variational Response Equilibrium

4.1 Introduction

This chapter discusses a Bayesian framework for games with incomplete information and possibly continuous action spaces. In particular, we define a framework in which each player's true payoff functions are unobserved by either the modeler or the other players. Each player's payoff function is assumed to be continuous with respect to the action space and is the sum of two terms: an observable component and an unobserved component, both of which are assumed to be continuous with respect to the action space.

For any game of complete information, Γ_0, with continuous payoff functions, we define a class of incomplete information games which possess the payoffs of Γ_0 as the observable component of players' payoffs. We restrict the incomplete information to be with respect to continuous variations of these payoffs.

Using this class of extensions of Γ_0, we define the set of (possibly mixed) strategy profiles which are rationalizable as Bayesian Nash equilibria of such an incomplete information extension of Γ_0, Γ. We term any element of this set a *variational response equilibrium* of Γ_0.

Our framework is similar to the work of several other scholars. Most recently, our work is closely related to the notion of quantal response equilibrium (QRE), first defined by McKelvey and Palfrey for both extensive-form (1998) and normal-form games (1995) with finite action spaces. A similar notion of equilibrium is due to Chen, Friedman, and Thisse (1997), though their notion of boundedly rational Nash equilibrium is narrower in scope than the QRE.

Earlier work on issues closely related to those examined here includes the work of Harsanyi (1967-68), (1973) on perturbed games and the work of Aumann, *et al.* (1983) on purification of mixed strategies. The work of Milgrom and Weber (1985) is a *seminal* contribution to the understanding of Bayesian equilibria in games of incomplete information. Milgrom and Weber define a distributional strategy as a measure over the Cartesian product of a player's action and type spaces. Our Theorem 58 rests upon their equilibrium existence result (Theorem 1 in Milgrom and Weber (1985)). Similarly, the logic behind our Theorem 59 closely resembles Milgrom and Weber's existence result as well.

This framework has many similarities to that examined in Milgrom and Weber (1985). On the technical side, the conditions of our Theorem 58 satisfies their Assumptions R1 (Equicontinuous payoffs) and R2 (Absolutely Continuous Information). Substantively, however, our framework differs in its motivation. Milgrom and Weber established existence results for a very general class of games of incomplete information. Our motivation is to provide a particular foundation for the study of games in which payoff perturbations are continuous with respect to the players' actions spaces. Thus, while we do provide a result (Theorem 59) using assumptions whose relationship to Milgrom and Weber's R1 and R2 is not clear,[1] the point of this chapter is essentially to direct attention at a specific application of their results.

This chapter provides a Bayesian framework for understanding behavior in strategic situations within a framework of unobserved payoff disturbances. In many cases, theorists assume that agents possess continuous payoff functions. This framework does not depart from this assumption. Our generalization of the perturbed games literature is to allow for utility perturbations which are correlated across possibly a continuum of pure actions. We allow this both for mathematical generality as well as possibly increased empirical realism. In addition, we can allow for correlation between the utility perturbations received by different players (Theorem 59). Our main result (Theorems 58), however, assumes that each player's shock is independent of her opponents' perturbations.

We first define the notion of *variational response equilibrium* (VRE). Next we prove existence of variational response equilibria as well as the existence of equilib-

[1] That is, the conditions under which Theorem 59 holds are not nested with Milgrom and Weber's Assumptions R1 and R2.

rium in a class of games of incomplete information which satisfy a weaker informational requirement. In addition, we provide some discussion of the similarities and differences between QRE in finite games and VRE in finite and continuous games.

Unfortunately, our definition does not appear to yield itself easily to empirical application, as our type space is the space of continuous functions on a metrizable action space. A different tact one could take in attempting to extend the definition would be to begin with the desired characteristics of individual behavior, such as that more costly mistakes are no more likely to be made than less costly ones, for example, and then examine the properties of fixed points of quantal response correspondences satisfying these characteristics.

We define our primitives in Section 4.2. Our existence results are found in Section 4.3. Conclusions are offered in Section 4.4.

4.2 The Model

In this section we describe the basic assumptions of our model. We consider normal-form games. Let N denote the set of players, with $|N| = n < \infty$.

4.2.1 Action Spaces

Let A_i denote the action space of player i. For all $i \in N$, we assume that A_i is a compact subset of a complete and separable metric space, and write $A = \prod_{i \in N} A_i$.

Assumption 48 *For all $i \in N$, A_i is a compact subset of a complete and separable metric space.*

4.2.2 Preferences and Types

Each player $i \in N$ is endowed with a payoff function $u_i : A \to \Re$. We assume that $u_i \in \mathcal{C}(A; \mathbb{R})$.

Assumption 49 *For all $i \in N$, $u_i \in \mathcal{C}(A; \mathbb{R})$.*

Each player $i \in N$ possesses a type, $\eta_i \in T_i = \mathcal{C}(A_i; \mathbb{R})$, which represents an additive payoff perturbation. The player receives the payoff perturbation associated with the action she chooses, regardless of the other players' actions. We write $T = \prod_{i=1}^{n} T_i$ for the space of all possible type profiles.

Formally, let $\mathcal{B}(X)$ denote the Borel σ-algebra on $\mathcal{C}(X; \mathbb{R})$, where X denotes any compact subset of a complete and separable metric space. For each $i \in N$, let ρ_i be an atomless probability measure defined on $\mathcal{B}(A_i)$.[2] As we state formally below, we will denote the resulting product measure on $\prod_{i \in N} \mathcal{B}(A_i)$ by ρ, the joint distribution of types, which is represented by an atomless measure since each ρ_i is assumed to be atomless.

Assumption 50 *The distribution of η is represented by a probability measure,*

$$\rho = \prod_{i \in N} \rho_i.$$

Where each ρ_i is atomless. Such a probability measure ρ is referred to as admissible.

Assumption 50 is equivalent to assuming that players' types, or utility perturbations, are independently distributed. As alluded to above, it is possible to prove existence of Nash equilibria in distributional strategies by assuming only that the joint distribution of types is atomless on T, which we show formally in Theorem 59.

In our framework, the payoff functions are assumed to be continuous with respect to the action space and the types enter in a simple and very particular fashion. As discussed above, and now defined formally, we assume that the privately observed types affect payoffs in an additively separable fashion.

Assumption 51 *Given an action profile $a \in A$, and type $\eta_i \in C_b(A_i)$, player i receives a payoff of*

$$v_i(a; \eta_i) = u_i(a) + \eta_i(a_i). \tag{4.1}$$

[2]Such a measure exists since A is compact and hence $\mathcal{C}(A; \mathbb{R})$ is Polish (Aliprantis and Border (1994), Theorem 11.58, p.407).

Let $U = \{u_1, \ldots, u_n\}$ and $v = \{v_1, \ldots, v_n\}$. Any game $\Gamma = (U, N, A, v, \rho)$ satisfying Assumptions 49-51 is termed *admissible*. For clarity, we denote the game of complete information corresponding to Γ by $\Gamma_0 = (U, N, A)$, and will often write $\Gamma = (\Gamma_0, v, \rho)$ when the context is clear.

4.2.3 Distributional Strategies

Distributional strategies are a means of describing mixtures over possibly uncountably infinite action and type spaces.

Definition 52 *A distributional strategy for player i is a probability measure μ_i : $A_i \times T_i \rightarrow [0, 1]$ where the marginal of μ_i on T_i is equal to ρ_i.*

The notion of a distributional strategy is meant to represent a possibly mixed strategy while avoiding measurability problems with continuous action and/or type spaces. The requirement that the marginal distribution of a players' distributional strategy with respect to her type equals the true distribution of her type represents the fact that a player can not change her type distribution.

4.2.4 Expected Payoffs

The expected payoff for player i of action $a_i \in A_i$, given type η_i and opponents' strategies μ_{-i}, is given by

$$V_i(a_i; \eta_i, \mu_{-i}) = E_{\mu_{-i}}[u_i(a_i; a_{-i}) + \eta_i(a_i)]$$

We denote the set of all distributional strategies for player i by M_i, and write $M = \prod_{i=1}^{n} M_i$ for the set of all possible vectors of distributional strategies. For any player $i \in N$, we write $\mu_{-i} = \{\mu_1, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_n\}$ for the profile of i's opponents' distributional strategies. The space of all such vectors for a given player i is denoted by M_{-i}.

We will denote the expected payoff of player i, given distributional strategy profile

μ, by $W_i(\mu)$. Formally,

$$W_i = E_\mu[u_i(a) + \eta_i(a_i].\tag{4.2}$$

The next result notes that W_i is linear in μ_i.

Lemma 53 *For $i \in N$, any $\alpha \in [0, 1]$, any $\beta \in \mathbb{R}$, and any $\mu_{-i} \in M_{-i}$, and any pair of distributional strategies $\mu_i, \mu_i' \in M_i$,*

$$W_i(\alpha\mu_i + (1-\alpha)\mu_i'; \mu_{-i}) = \alpha W(\mu_i; \mu_{-i}) + (1-\alpha)W(\mu_i'; \mu_{-i})$$

and

$$W_i(\beta\mu_i; \mu_{-i}) = \beta W(\mu_i; \mu_{-i})$$

Proof: The result follows from the fact that integration is a linear functional. ∎

4.2.5 Equilibrium in Distributional Strategies

Definition 54 *A Nash equilibrium in distributional strategies is an N-tuple, $\mu^* = (\mu_1^*, \ldots, \mu_n^*)$ for which, for all $i \in N$ and all $\hat{\mu}_i \in M_i$,*

$$W_i(\mu) = E_{\mu_i}[V_i(a_i; \eta_i)] \geq E_{\hat{\mu}_i}[V_i(a_i; \eta_i)\hat{\mu}_i] = W_i(\hat{\mu}_i; \mu_{-i}).$$

In words, a Nash equilibrium is a vector of distributional strategies at which no player has an alternative distributional strategy that would result in a strictly higher payoff.

We now define a ρ-Variational Response Equilibrium of a game of complete information, Γ_0, as any Nash equilibrium in distributional strategies of the game of incomplete information given by (Γ_0, T, ρ), where ρ is admissible, and $T = \mathcal{C}(A; \mathbb{R})$, endowed with its Borel σ-algebra.

Definition 55 *For an admissible game $\Gamma = (\Gamma_0, T, \rho)$, let μ^* be a Nash equilibrium in distributional strategies. Then the marginal distribution of μ^* with respect to the ac-*

tion space A, denoted by $\sigma_\rho^{\mu^*} : \mathcal{B}(A) \to [0,1]$ *is a* ρ-Variational Response Equilibrium *of* Γ_0.

Definition 56 *Let* Γ_0 *be a game of complete information and* $(\sigma_1, \ldots, \sigma_N)$ *be a vector of probability measures, with* $\sigma_i : \mathcal{B}(A_i) \to [0,1]$. *Then the product measure* $\sigma = \prod_{i=1}^N \sigma_i$ *is a* Variational Response Equilibrium *of* Γ_0 *if there exists an admissible probability measure* ρ, *defined on* $\mathcal{B}(\mathcal{C}(A;\mathbb{R}))$, *and a* ρ*-Variational Response Equilibrium* μ^* *satisfying*

$$\sigma(a) = \int_T \mu^*(a,t)dt$$

for all $a \in A$.

Thus, one can view the set of variational response equilibria for a game of complete information Γ as being parameterized by ρ, since the nature of v, i.e., that the payoff shocks are additive with respect to pure actions, is specified by the definition of admissibility.

4.3 Existence

In this section we prove existence of variational response equilibria for all games possessing continuous payoff functions. The next lemma states that, for a given ρ, the set of distributional strategy profiles is closed in the weak* topology. This result is referred to by Milgrom and Weber (1985) in the proof of their Theorem 1, but not shown. In addition, Mas-Colell (1984) implicitly refers to this result in the proof of his Theorem 1. Being unable to find a formal proof of this fact, it is included here.

Lemma 57 *Given a type distribution* ρ, *the set of distributional strategy profiles,* $M(A \times T; \rho)$, *is a closed subset of* $\mathcal{P}(A \times T)$ *when endowed with the weak* topology.*

Proof: When endowed with the relative weak* topology induced by the weak* topology on $\mathcal{P}(A \times T)$, $M(A \times T; \rho)$ is a metric space. Thus, one can verify its closure by checking sequences. In addition, the separability and completeness of A and T ensure that $\mathcal{P}(A \times T)$ is complete and separable (i.e., Polish) when endowed with the weak* topology (Aliprantis and Border (1994), Theorem 14.15).

We know that $M(A \times T; \rho)$ is a tight set of probability measures[3] because $A \times T$ is Polish, and every finite Borel measure on a Polish space is regular (Aliprantis and Border, (1994), Theorem 10.7).

By Prohorov's theorem (Billingsley (1968), p. 240), M is a relatively compact subset of $\mathcal{P}(A \times T)$ when endowed with the weak* topology.[4] Thus, let $\{\mu_\alpha\}$ be a sequence in $M_\rho(A \times T)$. The relative compactness of M ensures that $\{\mu_\alpha\}$ contains a weak* convergent subsequence, so we assume without loss of generality that $\{\mu_\alpha\}$ is itself a weak* convergent sequence, with limit μ necessarily in $\bar{M}$, the closure of M.

We now show that the limit, μ, is itself a distributional strategy profile in $M(A \times T; \rho)$. To show this, we must show three things: (1) $\mu(A \times T) = 1$, (2) μ is nonnegative for all Borel subsets of $A \times T$, and (3) the marginal of μ with respect to T equals ρ. We now proceed to show these in order.

(1) From Theorem 14.3 in Aliprantis and Border (1994), weak* convergence of μ_α to μ is equivalent to

$$\limsup_\alpha \mu_\alpha(F) \leq \mu(F) \tag{4.3}$$

for each closed set $F \in A \times T$, and

$$\liminf_\alpha \mu_\alpha(G) \geq \mu(G) \tag{4.4}$$

for each open set $G \in A \times T$.

Letting $F = A \times T$ (which is a closed set since the empty set is open), it is trivial to see that $\limsup_\alpha \mu_\alpha(A \times T) = \limsup\{1, 1, \dots ,\} = 1$. Then, if $\mu_\alpha \Rightarrow \mu$, it must be the case that $\mu(A \times T) \geq 1$.

---

[3] A set $\mathcal{F}$ of probability measures on a space X is *tight* if, for each $\varepsilon > 0$ there exists a compact set K satisfying $\mu(K) > 1 - \varepsilon$ for each $\mu \in \mathcal{F}$.

[4] A subset Y of a topological space X is *relatively compact* if its closure is compact.

Similarly, note that F is also an open Borel subset of $A \times T$, and $\liminf_\alpha \{1, 1, \ldots, \} = 1$, so that $\mu_\alpha \Rightarrow \mu$, implies $\mu(A \times T) \leq 1$. Therefore, $\mu_\alpha \Rightarrow \mu$ implies that $\mu(A \times T) \geq 1$ and $\mu(A \times T) \leq 1$, meaning that $\mu(A \times T) = 1$.

(2) Since $\mu_\alpha \in M(A \times T; \rho)$ for all α, any closed set $B \subset A \times T$, $\limsup_\alpha \mu_\alpha(B) \geq 0$. By Equation 4.3,

$$\begin{aligned} \mu(B) &\geq \limsup_\alpha \mu_\alpha(B) \\ &\geq 0, \end{aligned}$$

implying that μ is nonnegative for all closed subsets of $A \times T$.

(3) For all α,

$$\int_A d\mu_\alpha(a, \eta) = \rho.$$

Since integration is a linear functional, for any sequence μ_α it must be the case that

$$\int_A \lim_{\alpha \to \infty} d\mu_\alpha(a, \eta) = \lim_{\alpha \to \infty} \int_A d\mu_\alpha(a, \eta),$$

implying that

$$\int_A d\mu(a, \eta) = \rho.$$

Thus, the marginal of μ with respect to T is equal to ρ. Therefore, $\lim_{\alpha \to \infty} \mu_\alpha \in M(A \times T; \rho)$ for all convergent sequences $\{\mu_\alpha\}$, so that M is closed, completing the proof. ∎

Let $\phi_i : M_{-i} \to\to M_i$ denote the best response correspondence for player i and $\phi = \prod_{i=1}^n \phi_i$.

Theorem 58 *Let Γ be admissible, and $u_i \in \mathcal{C}(A; \mathbb{R})$ for all $i \in N$. Then Γ_0 possesses a Variational Response Equilibrium.*

Proof: We show that conditions R1 and R2 for Theorem 1 in Milgrom and Weber (1985) are satisfied. Endowing T with the topology of uniform convergence, $v_i : T \times A \to \mathbb{R}$ is a uniformly continuous function. Hence, by Proposition 1 in Milgrom and Weber (1985), Condition R1 is satisfied. The random variables $\{\tau_1, \ldots, \tau_N\}$ are mutually independent so that, by Proposition 3 in Milgrom and Weber (1985), Assumption R2 is satisfied. Therefore, by Theorem 1 in Milgrom and Weber (1985), there exists an equilibrium in distributional strategies. The projection of such an equilibrium onto A is a variational response equilibrium of Γ_0. ∎

We now digress for a moment to prove an existence result under the assumption that the joint distribution of types, ρ, possesses full support and is atomless. Only the latter of these conditions is implied by our earlier assumptions. One implication of our requirement is that the players can assign no profile of the other players' types zero probability conditional upon their own type. Nevertheless, their conditional beliefs about the other players' types may differ according to the realization of their own type, of course.

Theorem 59 *Let ρ be an atomless distribution on $T = \prod_{i=1}^{n} \mathcal{C}(A_i)$ and $u_i : A \to \mathbb{R}$ be continuous for each $i \in N$. Then Γ_0 possesses a Variational Response Equilibrium.*

Proof: For all $i \in N$, W_i is an integral of the sum of continuous functions,

$$W_i(\mu) = \int_{T \times A} [u_i(a) + \eta_i(a_i)]\, d\mu(a, \eta).$$

Endowing M with the product weak* topology, W_i is continuous on M. Since T is a complete and separable metric space, ρ is a tight probability measure.

Since ρ is tight, it follows from Prohorov's Theorem that M is a relatively compact set for each $i \in N$ (see Billingsley (1968), p. 240). By Lemma 57, M is closed. The closure of a relatively compact set is compact, so M is compact.

By the continuity of W_i and Berge's Theorem of the Maximum (Aliprantis and Border (1994), Theorem 14.30), the graph of ϕ_i is closed for each $i \in N$. Every closed subset of a compact topological space is compact, so that the graph of ϕ_i is compact.

In addition, the graph of ϕ_i is convex by the linearity of W_i for all $i \in N$ (Lemma 53). In addition, having closed graph implies that ϕ_i is upper hemicontinuous since M_i is a compact space. The product of upper hemicontinuous correspondences is upper hemicontinuous (see Border, (1985), Proposition 11.25). This implies that ϕ is an upper hemicontinuous correspondence.

Thus, by Fan's fixed point theorem (1952), there exists $\mu \in M$ such that $\mu \in \phi(\mu)$. Such a fixed point is a Nash equilibrium in distributional strategies, completing the proof. ∎

4.4 Comparisons, Extensions, and Conclusions

In this chapter we have defined and explored the notion of variational response equilibrium. The notion of VRE is in some senses a generalization of the notion of quantal response equilibrium in finite games. It is also related to the more general literature on games with perturbed payoffs. Primarily, our definition of VRE allows for correlation of payoff perturbations across actions. Indeed, correlation of payoff disturbances across actions is required within our framework (except as a limiting case) whenever the action space is continuous. In addition, we have shown that certain types of correlation can be allowed between the realizations of players' types (Theorem 59).

Several questions follow from our definition. For example, are any probability measures on A *not* variational response equilibria of Γ_0? That is, do there exist mixed strategy profiles that can not be rationalized according to an admissible game of incomplete information? One conjecture along this dimension is that any measure, $\sigma : \mathcal{B}(A) \to [0, 1]$ induced by a VRE must have full support on A. This is a characteristic of quantal response equilibria in finite games, as well as the notion of logit equilibrium used by Anderson, Goeree, and Holt. The validity of the conjecture is not known at this point, however. A related question is whether we can use this refinement to select certain equilibria of the complete information game, Γ_0.

4.4.1 VRE and QRE in Finite Games

The notion of quantal response equilibrium, as discussed earlier, has been defined as a Bayes Nash equilibrium in games with finite action spaces. Our definition of variational response equilibrium is also defined for such games, given a metric on the action space, A. This section discusses the relationship between QRE and VRE in finite normal form games.

The idea underlying quantal response equilibrium is that players are unable to commit to playing any action with zero probability. This lack of commitment ability may be due to any of several factors, including imperfect implementation of a pure strategy (e.g., trembles), incorrect beliefs, or idiosyncratic shocks to each player's preferences. Thus, the notion of QRE is both a theoretical and an empirical tool. Theoretically, the QRE framework provides an environment in which to examine the properties of games in which there is always a positive probability that each action profile will be observed. On the other hand, the QRE has already been fruitfully applied in attempts to explain actual behavior in the laboratory (see McKelvey and Palfrey (1995), (1998)).

McKelvey and Palfrey define the QRE as a Bayesian equilibrium of a game in which each player observes independently and identically distributed shocks to the expected payoff for each action available to her. Thus, the set of quantal response equilibria is parameterized by the distribution from which these payoff perturbations are drawn. This formulation implies that payers are more likely to choose actions with higher expected payoff, ceteris paribus.

The empirical motivations behind VRE are similar in spirit to those behind QRE. Both equilibrium concepts are motivated by a desire to provide stochastic explanations for deviations from Nash equilibrium play. The substances of the two notions are not identical, however. In particular, even in finite games, the notion of variational response equilibrium is sensitive to the "distance" between two actions. This is because, for any player in a finite game, the distance between actions determines the correlations between that player's action-specific utility perturbations. In a QRE, *each action's utility* perturbation is realized independently, regardless of any metric structure which the action space may possess.

For finite games in which some player has more than two pure actions, the type distributions which are admissible in our framework contain those distributions which McKelvey and Palfrey (1995) classify as admissible for quantal response equilibria as a strict subset. In fact, we can provide the following nesting result, which states that, for every finite normal form game Γ_0, if σ^* is a probability distribution on A which is induced by some quantal response equilibrium with payoff perturbations distributed according to an admissible probability measure F, then there exists an admissible type distribution ρ and type space T which generate a variational response equilibrium which induces σ on A as well.

Let A_i be finite for every player i, with $|A_i| = \alpha_i$, and let $\Gamma_0 = (N, A, U)$. For each $i \in N$, relabelling A_i, define $\hat{A}_i = \{1, \ldots, \alpha_i\}$. Now define T_i to be set of piecewise continuous functions from $[1, \alpha_i]$ defined as

$$T_i = \{g \in \mathcal{C}([0, \alpha_i]; \mathbb{R}) : \forall r \notin \hat{A}_i, g(r) = (g(\lfloor r \rfloor) - g(\lceil r \rceil))(r - \lfloor r \rfloor)\},$$

and

$$\rho_i = \prod_{j=1}^{\alpha_i} \rho_i^j,$$

with

$$\rho_i^j(r) = \Pr[g(j) \leq r] = F_i(r),$$

for all $j \in \hat{A}_i$, where F_i is the cumulative density function of the payoff perturbations which generates σ as a quantal response equilibrium of Γ_0 (see McKelvey and Palfrey (1995)). Thus, by enriching the type space and choosing the type distribution judiciously, any QRE of a finite normal form game can be justified as a variational response equilibrium. The converse is easily seen to be false.

There are other important differences aside from the fact that every QRE is also a *VRE in finite* games. In particular, if the game being played is a multistage game, even in normal form, the choice of metric over the space of all pure actions for any given player may have an enormous impact on the set of variational response equilibria of such a game, due to the presence of correlation in the unobserved payoff shocks.

We feel that the notion of QRE in finite games, as defined by McKelvey and Palfrey, is best thought of as either representing the unobservable characteristics of players in a game or as a statistical model of bounded rationality. The latter interpretation is qualified by the restriction that players are less likely to make a mistake the more costly that mistake is in expectation. We have presented a notion, that of variational response equilibrium, which is similar in spirit and which we feel can be motivated in the same fashion.

One technical difference between the notions is that the VRE is defined for games with possibly continuous action spaces. But two key elements of our framework, regardless of the action space, are that variational response equilibrium requires that the action space be endowed with a metric, and that the payoff shocks observed by any given player are correlated with respect to this metric. Indeed, for the case of continuous action spaces, we suspect that this correlation is necessary for payoff maximization to make any sense. Regardless, this metric is an exogenous (and seemingly unobservable) parameter of the framework.

4.4.2 VRE: A Continuous Version of QRE?

One might ask whether variational response equilibrium is the natural extension of quantal response equilibrium to games with continuous action spaces. We argue that it is not, for the following reasons. First, the comparative statics of the QRE in finite games are similar to many models of probabilistic choice, such as the Luce model (1959) and the random payoff maximization framework of McFadden (1981), among others. We discuss below why such comparative statics, such as the fact that the probability of a player choosing some strategy x is weakly increasing in the expected payoff offered by x, may not hold for VRE.

Second, at least one of the parameterizations of the QRE, the logit QRE, has an *appealing functional* analogue in the continuous case, termed by Anderson, Goeree, and Holt (1998) the "logit equilibrium."

Let Γ be an n player game. In a logit equilibrium, for each player $i \in N$, the probability that an action x with corresponding expected payoff $U_i(x)$ is chosen by player i is given by

$$f_i(x) = \frac{e^{U_i(x)}}{\int_{A_i} e^{U_i(x)}}. \tag{4.5}$$

This functional form has been used both theoretically and empirically by several researchers, including Anderson, Goeree, and Holt (1997), (1998), and Capra, Goeree, Gomez, and Holt (1997). We discuss below the potential problems with rationalizing the logit equilibrium as a variational response equilibrium.

Finally, a central assumption of VRE is that there is an ordering of each player's action space. In particular, each player's action space is endowed with a metric. The perturbation associated with an action is more highly correlated with the perturbations of nearby actions than with those of actions further away. In a quantal response equilibrium, each player's choice probabilities are invariant to a "shuffling" of the strategies - the ordering of the strategies is irrelevant. We show below that this is not the case in general for variational response equilibria.

Comparative Statics

A key qualitative element of Quantal Response Equilibria in finite games is that actions with higher expected payoff (given a distributional strategy profile for player i's opponents) are played with higher probability and choice probabilities are continuous functions of the expected payoffs of the strategies. Variational response equilibria may not have the continuity property, and very likely do not possess the monotonicity property in general, as two actions may yield the same expected payoff but have very different expected payoffs due to the expected payoffs of other actions which are "close" to either of them.

Functional Forms

The fact that comparative statics are difficult to provide for general variational response equilibria highlights a more fundamental problem: what do variational response equilibria look like? That is, one would like to be able, even if only for a restricted class of type distributions, to write down a function of the observable payoffs which generates a set of variational response equilibria to a given game of complete information, Γ_0.[5]

It seems that studying random payoff maximization in an environment with a possibly uncountable choice set is not only difficult, but that the links that exist between functional form models, such as that used by Anderson, Goeree, and Holt (1997), (1998), and Capra, Goeree, Gomez, and Holt (1997), and models based in a Bayesian equilibrium setting, such as that defined here, are not at all clear. As an example, it is not known at this time whether or not there exists any admissible type distribution ρ for which the notion of *logit equilibrium* employed in Anderson, Goeree, and Holt (1998) is a ρ-variational response equilibrium. It seems that the answer to this question is no, since the logit equilibrium shares the monotonicity property possessed by the quantal response equilibrium as discussed above. In particular, for any player $i \in N$ any two actions x and y in A_i satisfying $U_i(x) = U_i(y)$, $f(x) = f(y)$, meaning that, conditional on choosing either x or y, each is equally likely to be chosen in a logit equilibrium. A more rigorous answer to this question is left as a topic for future research.

The Action Space

As stated earlier, the notion of variational response equilibrium imposes both a metric on the action space as well as a framework of unobserved payoff shocks, while the notion of quantal response equilibrium only imposes a framework of unobserved payoff shocks. Each player's payoff shocks in the variational response equilibrium

[5]Indeed, as Tom Palfrey has pointed out to me, the notion of quantal response equilibrium began as a functional form; in particular, as the discrete version of the logit equilibrium discussed above. See McKelvey and Palfrey (1996).

are *correlated* with respect to the metric imposed on her action space. This is very different from the assumption of *independent* payoff shocks which rationalize quantal response equilibria from a Bayesian perspective.

4.4.3 Extensions

Possible extensions of this framework include allowing for the payoff perturbation to enter in ways other than as an additive shock. In addition, there are many alternative type spaces which could be examined. We have explored the possibility of directly extending the definition employed by McKelvey and Palfrey (1995) by assuming that each player is endowed with a continuum of independently and identically distributed random perturbations, one for each pure action. It turns out that the arg max of the resulting v_i function is not even guaranteed to be Lebesgue measurable. This problem is related in an intimate fashion to the use of the Axiom of Choice which is required by the uncountable nature of the action space. Thus, such an extension does not seem to even make sense, much less be tractable.

A more promising extension would be to take a space of random walks (which is, of course, a subset of the type space assumed here) generated by a particular random process, such as a normally distributed "step" (i.e., Brownian motion). The potential value of such an approach largely consists of the ability to say more about the characteristics of behavior in such a framework. In addition, such a definition of the type space may be more appealing from a descriptive standpoint.

In short, much remains to be developed in the theory of perturbed games. We argue that any complete theory must begin to place a structure on the game being played. By placing a structure we mean that the theory must account for explicitly behavioral effects of the game's context, design, and representation.

We have attempted to start this endeavor by placing a metric on the action space, a technique which is a very preliminary attempt at capturing, for example, the similarity of actions as perceived by the players. Other aspects of games which a more complete theory might take into account include the implications of different players' roles (e.g., whether a player perceives some of his opponents as having more or *less of an advantage*, whether a player infers something about other players' future behavior from his or her own perecptions, etc.), differences in mental representations of stochastic processess, the effect of the order in which decisions are made in an extensive-form game, and the representation of payoffs (e.g., payoffs as losses versus gains, payoffs expressed relative to other players, etc.), among what is surely a large number of other possibilities.

Bibliography

Aliprantis, Charalambos D. and Kim C. Border. 1994. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Berlin: Springer-Verlag.

Anderson, Robert M. and William R. Zame. 2000. "Genericity with Infinitely Many Parameters." *Mimeo*, University of California, Los Angeles.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt. 1997. " Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction." *Mimeo*, University of Virginia.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt. 1998. "Minimum-Effort Coordination Games." *Mimeo*, University of Virginia.

Anderson, S.P., Katz, A. and Thisse, J.-F. (1992). "Probabilistic Voting and Platform Selection in Multi-Party Elections." Unpublished.

Aranson, P. H., M. J. Hinich, and P. C. Ordeshook. 1974. "Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives." *American Political Science Review* 68: 135–52.

Aumann, Robert J., Y. Katznelson, R. Radner. R. W. Rosenthal, and B. Weiss. 1983. "Approximate Purification of Mixed Strategies." *Mathematics of Operations Research* 8:327-341.

Banks, Jeffrey S. and John Duggan. 1999. "The Theory of Probabilistic Voting in the Spatial Model of Elections." *Mimeo*, California Institute of Technology.

Baron, David P. and John A. Ferejohn. 1989. "Bargaining in Legislatures." *American Political Science Review* 83(4):1181–1206.

Bhattacharya, R. N. and R. R. Rao, *Normal Approximation and Asymptotic Expansions*. (Malabar, FA: R. E. Krieger Publishing Co, 1986)

Billingsley, P. 1968. *Convergence of Probability Measures.* New York: John Wiley and Sons.

Black, Duncan. 1958. *The Theory of Committees and Elections.* London: Cambridge University Press.

Block, H. D. and J. Marschak. 1960. "Random Orderings and Stochastic Theories of Response," in *Contributions to Probability and Statistics* (I. Olkin, Ed.) pp. 97-132. Stanford: Stanford University Press.

Border, Kim C. 1985. *Fixed Point Theorems with Applications to Economics and Game Theory.* Cambridge: Cambridge University Press.

Capra, C. Monica, Jacob K. Goeree, Rosario Gomez, and Charles A. Holt. 1997. "Anomalous Behavior in a Traveler's Dilemma?" *Mimeo*, University of Virginia.

Chen, H.-C., J. W. Friedman, and J.-F. Thisse. 1997. "Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach." *Games and Economic Behavior* 18: 32-54.

Coughlin, P. J. 1983. "Social Utility Functions for Strategic Decisions in Probabilistic Voting Models," *Mathematical Social Sciences* 4: 275-293.

Coughlin, P. J. 1992. *Probabilistic Voting Theory.* Cambridge: Cambridge University Press.

Coughlin, P. J., and S. Nitzan. 1981. "Electoral Outcomes with Probablilistic Voting and Nash Social Welfare Maxima." *Journal of Public Economics* 15: 113-122.

Coughlin, P. J. and S. Nitzan. 1981. "Directional and Local Electoral Equilibria with Probabilistic Voting." *Journal of Economic Theory* 24: 226-40.

Dasgupta, Partha and Eric Maskin. 1986. "The Existence of Equilibrium in Discontinuous Games, I: Theory." *Review of Economic Studies* 53: 1-26.

dePalma, A., Hong, G. S., and Thisse, J. F. 1990. "Equilibria in multi party competition under uncertainty." *Social Choice and Welfare* 7: 247-259.

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

Duggan, John. 2000. "Equilibrium Equivalence Under Expected Plurality and Probability of Winning Maximization." *Mimeo*, University of Rochester.

Eaton, B. C. and R. G. Lipsey. 1975. "The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition." *Review of Economic Studies* 42: 27-50.

Edgell, S. E. and W. S. Geisler. 1980. "A Set-Theoretic Random Utility Model of Choice Behavior." *Journal of Mathematical Psychology* 21:265-278.

Enelow, J. M. and M. J. Hinich. 1989. "A General Probabilistic Spatial Theory of Elections." *Public Choice* 54: 101-113.

Enelow, J. M. and M. J. Hinich. 1990. *Advances in the Spatial Theory of Voting.* Cambridge: Cambridge University Press.

Fan, Ky. 1952. "Fixed-Point and Minimax Theorems in Locally Convex Topological Linear Spaces." *Proc. Nat. Acad. Sci. U.S.A.* 38:121-26.

Glicksberg, I.L. 1952. "A Further Generalization of the Kakutani Fixed Point Theorem, with Applications to Nash Equilibrium Points." *Proc. Amer. Math. Soc.* 3:170-174.

Harsanyi, J.C. 1967-68. "Games with Incomplete Information Played by Bayesian Players." *Management Science* 14: 159-82, 320-334, 486-502.

Harsanyi, J.C. 1973. "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed Strategy Equilibrium Points." *International Journal of Game Theory* 2: 1-23.

Hinich, M. 1977. "Equilibrium in Spatial Voting: The Median Voter Result is an Artifact." *Journal of Economic Theory* 16: 208-219.

Hinich, M., J. Ledyard, and P. Ordeshook. 1972. "Nonvoting and the Existence of Equilibrium under Majority Rule." *Journal of Economic Theory* 4: 144-153.

Hinich, M. J. and P. C. Ordeshook. 1972. "Plurality Maximization vs. Vote Maximization: A Spatial Analysis with Variable Participation." *American Political Science Review* 64: 772–91.

Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal* 39:41-57.

Hunt, Brian R., Tim Sauer, and James A. Yorke. 1992. "Prevalence: A Translation-Invariant 'Almost Every' on Infinite-Dimensional Spaces." *Bulletin (New Series) of the American Mathematical Society* 27: 217-238.

Ledyard, J. 1984. "The Pure Theory of Large Two-Candidate Elections." *Public Choice* Vol. 44: 7-41.

Lin, T. M., J. M. Enelow, and H. Dorussen. 1999. "Equilibrium in multicandidate probabilistic voting," *Public Choice* 98: 59-82.

Luce, D. 1959. *Individual Choice Behavior*. New York: Wesley.

Mas-Colell, Andreu. 1984. "On a Theorem of Schmeidler." *Journal of Mathematical Economics* 13: 201-206.

McFadden, D. 1981. "Econometric Models of Probabilistic Choice", in *Structural Analysis of Discrete Data with Econometric Applications* (C. F. Manski and D. McFadden, eds.) pp. 198-272. Cambridge: MIT Press.

McKelvey, R. D. and P. C. Ordeshook. 1972. "A General Theory of the Calculus of Voting," in *Mathematical Applications in Political Science VI* (J. Herndon and J. Bernd, eds.) pp. 32-78. Charlottesville: University Press of Virginia.

McKelvey, R. D. and T. R. Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *Japanese Economic Review* 42: 186-209.

McKelvey, R. D. and T. R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10: 6-38.

McKelvey, R. D. and T. R. Palfrey. 1998. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics* 1: 9-41.

McKelvey, R. D. and J. W. Patty. 1999. "A Theory of Voting in Large Elections." California Institute of Technology Social Science Working Paper 1056.

Milgrom, P., and R. Weber. 1985. "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research* 10: 619-632

Munroe, M.E. 1971. *Measure and Integration.* Reading, Massachusetts: Addison-Wesley.

Myerson, R. 1997. "Large Poisson Games." MEDS Discussion paper # 1189, Northwestern University.

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

Nash, John F. 1950. "Equilibrium Points in N-Person Games." *Proceedings of the National Academy of Sciences, U.S.A.* 36: 48-49.

Patty, John W. 1999. "Plurality and Probability of Victory: Some Equivalence Results." California Institute of Technology Social Sciences Working Paper 1048.

Patty, J. W. 1999. "Equilibrium Equivalence with J Candidates and N Voters." California Institute of Technology Social Science Working Paper 1068.

Plott, C. R. 1967. "A Notion of Equilibrium and Its Possibility Under Majority Rule." *American Economic Review* 57: 787-806.

Rubenstein, A. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50: 97-110.

Serfling, R. 1980. *Approximation Theorems of Mathematical Statistics.* New York; Wiley.

Simon, Leo. 1987. "Games with Discontinuous Payoffs." *Review of Economic Studies* 54: 569-597.

Sion, M. and P. Wolfe. 1957. "On a Game Without a Value." *Contributions to the Theory of Games, III.* Princeton: Annals of Mathematical Studies No. 39, pp. 299-306.

Stokey, Nancy L. and Robert E. Lucas, Jr. with Edward C. Prescott. 1989. *Recursive Methods in Economic Dynamics.* Cambridge: Harvard University Press.

Taylor, Angus E. 1965. *General Theory of Functions and Integration.* New York: Blaisdell.

Thurstone, L. L. 1927. "A Law of Comparative Judgment." *Psychological Review* 34: 273-286.

Ville, J. 1938. "Sur la Théorie Générale des Jeux où Intervient l'habilité des Joueurs." *Traité de Calcul des Probabilités et de ses Applications*: E. Borel, et al, Ed. 2(5):105-113.

von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior.* Princeton: Princeton University Press. Second Ed., 1947.

Wang, Y. H. 1993. "On the Number of Successes in Independent Trials." *Statistica Sinica* 3: 295-312.

Wittman, D. A. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *American Political Science Review* 77: 142-57.