We propose an equilibrium for \( n \)-person finite games based on bounded rationality using the logit model of discrete choice theory. At equilibrium, each player uses appropriate choice probabilities, given those used by the others. Rationality is parameterized on a continuum from complete rationality to uniform random choice. Results on the existence of equilibrium and on convergence to Nash as rationality becomes perfect are similar to results due to McKelvey and Palfrey. We identify conditions such that for a given rationality parameter range the path of choices over time when the players use fictitious play converges to equilibrium. *Journal of Economic Literature* Classification Numbers: C72, L20.

© 1997 Academic Press
1. INTRODUCTION

Boundedly rational decision, that is, decisions by agents who imperfectly optimize, has been of interest for nearly half a century as evidenced by such work as Alchian (1950), Simon (1957), Winter (1964), and Nelson and Winter (1982). The great recent interest in and explosion of research in evolutionary game theory is also motivated to a significant extent by a desire to explain behavior by agents who do not optimize (see the November 1995 special issue of this journal). In much of the work cited above, agents behave in a way that tends toward choosing better alternatives more frequently than inferior ones or that tends to change choices over time in directions that appear to be beneficial (utility increasing). Thus the decision makers below follow behavior that exhibits a tendency toward utility maximization, but is not optimal. Agents' choices are made according to the random choice interpretation of discrete choice theory so that boundedly rational behavior is naturally expressed here by mixed strategies. No agent is consciously aware of a complete and transitive preference ordering of the sort conventionally assumed in economics. It is as if he has such a deterministic preference ordering buried rather deeply in his unconscious mind and his decisions reflect a rough recognition of this ordering. If one alternative is superior to another according to his subconscious preferences, it will have a higher probability of being chosen. The comparative values of these probabilities express the individual's tendency toward utility maximization.

In an important, interesting paper, McKelvey and Palfrey (1995) propose a model of stochastic choice in finite games. Our model can be viewed as a reinterpretation of theirs in terms of bounded rationality that we believe to be particularly useful. We also extend their results to show conditions under which repeated play leads to convergence to equilibrium and we provide an interpretation of their model that we believe to be particularly useful. The underlying decision structure used by McKelvey and Palfrey is well known in econometrics (see McFadden, 1981); however, variants of the mathematical structure have a long history in economics and psychology (see Thurstone, 1927; Luce, 1959; Block and Marschak, 1960; and Edgell and Geisler, 1980). Various interpretations can be given to explain the apparently stochastic choices made by agents. These include that (i) agents are conventional utility maximizers whose perception of the utility of their possible choices is subject to noise (ii) agents are conventional utility maximizers whose utility functions are randomly drawn from some specific family, and (iii) agents are not utility maximizers, but, instead choose randomly in a fashion that is influenced by a subconscious (or latent) utility function. In brief, these interpretations can be called (i) noise, (ii) random utility, and (iii) bounded rationality.
McKelvey and Palfrey (1995) deal mainly with the mathematical structure of choice; however, when they interpret the underlying decision processes and knowledge of the agents, they lean mainly toward the noise interpretation. Indeed, the noise and random utility interpretations are very close in the sense that, in both cases, the decision maker has some randomly drawn evaluations with respect to which she optimizes. The bounded rationality interpretation takes a fundamentally different view of the choice process. To illustrate, suppose an individual can choose any one of several consumption bundles \( x \in X \). Each bundle has a well-defined utility \( u(x) \) in the sense that if the person actually consumes \( x^* \) then the satisfaction she experiences is measured by \( u(x^*) \); however, this utility function is not explicitly known to her. In our framework it is latent, which is why we refer to it as subconscious utility, and it influences the probability with which she will choose \( x^* \).

Any mathematical structure commonly used with the noise or random utility interpretations can also be interpreted as a model of bounded rationality. The logit formulation used by McKelvey and Palfrey—and variants of that formulation—is parameterized in a way that is very appealing for a bounded rationality interpretation: as a certain parameter varies from zero to infinity, the choice behavior of the agent varies from placing equal probability on all alternatives, irrespective of their subconscious utility, to fully rational utility maximizing choice behavior. In the noise interpretation this same parameter measures the randomness in the agent's utility function. McKelvey and Palfrey prove existence of their stochastic choice equilibrium (which they call a quan
tal response equilibrium or QRE) and they show that, as the parameter alluded to above goes to infinity, the corresponding sequence of equilibria converge to a Nash equilibrium. In our bounded rationality interpretation we call their QRE a boundedly rational Nash equilibrium.

In any game the utility that a player obtains when playing a particular strategy is a function of the strategy choices of the other players. We suppose a population of players who repeatedly play the same game and assume that the subconscious utility that a player attaches to a given strategy depends on that player's subconscious belief about the strategies the other players will select. These beliefs, in turn, are assumed to be given by the observed choices that the other players made in the past. When a player uses a mixed strategy, then we make the natural assumption that other players observe the realization of the mixed strategy and not the mixed strategy itself. Thus we model a dynamic process using fictitious play (Brown, 1951; and Robinson, 1951) under which players' beliefs concerning the other players' choice probabilities are given by observed past behavior and we prove that, under certain conditions both beliefs and actual choices converge over time to equilibrium.
Section 2 develops the basic model for random choice logit players, indicating how individuals make strategy choices given the basic game model. Then the concept of boundedly rational Nash equilibrium is introduced in Section 3, where results from McKelvey and Palfrey (1995) are reviewed. In Section 4 we model an adjustment process in which the players begin with arbitrary beliefs about the environment within which they choose. Over time they update their beliefs on the basis of the actual past choices of other players and, eventually, their beliefs and their choices converge in probability to those consistent with a boundedly rational Nash equilibrium. Finally, Section 5 contains concluding remarks.

2. THE SINGLE PERIOD MODEL

In this section the basic elements of the single shot game are described and then random choice behavior as developed in discrete choice theory is introduced. We use below a particular random choice formulation that differs from that of McKelvey and Palfrey (1995); however, their formulation or many others could be used as well. The differences have definite consequences for equilibrium, but basic results on existence and on convergence as the rationality parameters go to infinity will still hold.

Let \((N, A, u)\) be a finite game defined as follows. The set of players is \(N = \{1, \ldots, n\}\). Each player \(i \in N\) has a set of pure strategies \(A^i = \{1, \ldots, m\}\) with elements \(a^i\), the set of pure strategy profiles is \(A = \times_{i \in N} A^i\) with elements \(a\), and \(A^{-i} = \times_{j \neq i} A^j\) is the set of strategy profiles of all players other than \(i\), with elements \(a^{-i}\). The subconscious utility of player \(i\) associated with the strategy profile \(a \in A\) is \(u^i_a\); \(u^i = (u^i_1, \ldots, u^i_n)\) and \(u = \times_{i \in N} u^i\). Each player knows \(N\) and \(A\); that is, each player knows who is in the game and the strategy sets available to each. However, each player is ignorant of all other players' subconscious utility functions and each has only that vague awareness of his own subconscious utility which accords with the random choice interpretation of discrete choice theory. Whether any of this is common knowledge is unimportant, because it will be seen that the style of play considered in this paper is independent of this condition.

For simplicity in exposition we assume that the decisions of each player are made according to the multinomial logit model (see McFadden, 1981) or, to be more precise, a slightly modified version of this model. For a single decision maker the logit model applies to choice over a finite set of alternatives \(X = \{x_1, \ldots, x_m\}\). If the subconscious utility of \(x_i\) is \(u_i \geq 0\) and if \(\sum_i u_i > 0\), then under the logit, the probability of choosing \(x_i\) is \(u_i^\mu/\sum_j u_j^\mu\) where \(\mu\) is a nonnegative parameter that measures the degree of rationality of the decision maker. As \(\mu\) goes to infinity, the choice
probability of the decision maker tends to one for the alternative having the highest subconscious utility and tends to zero for the others. If the alternative with the highest subconscious utility is not unique, then the probabilities of these alternatives sum to one and are equal. Thus, as \( \mu \to \infty \) behavior converges to full rationality. At the other extreme, as \( \mu \to 0 \) the choice probabilities become equal. Preferences do not affect choice probabilities. When \( \mu = 1 \) we have the Luce (1959) model of probabilistic choice. Note that choice probabilities are invariant up to transformation of the subconscious utility function by a positive multiplicative constant and are strictly positive as long as the \( u_j \) are strictly positive and \( \mu \) is finite.

Since the logit model is formulated for individual decision, not decisions in a game-theoretic context, it is necessary to specify what such behavior means in an interactive environment. We envision each player having a probability distribution over the choices available to the others that can be expressed as a probability distribution \( P^i \) over the elements of \( A^{-i} \) where \( P^i(a^{-i}) \) is the probability associated with \( a^{-i} \in A^{-i} \). We assume an expected subconscious utility, so that the subconscious utility of the \( j \)th pure strategy of player \( i \), given \( P^i \), is

\[
U^i_j(P^i) = \sum_{a^{-i} \in A^{-i}} P^i(a^{-i}) u^i_{j,a^{-i}}.
\]  

(1)

Although we use expected utility in Eq. (1) there appears to be no reason why a nonexpected utility formulation could not be substituted as long as \( U^i_j(P^i) \) were required to be continuous in \( P^i \).

Using the logit framework, the choice probabilities over pure strategies of player \( i \) are defined as

\[
p^i_j(P^i) = \frac{[U^i_j(P^i)]^{\mu_i}}{\sum_{k=1}^m [U^i_k(P^i)]^{\mu_i}},
\]  

(2)

where \( \mu_i \) is the rationality parameter of player \( i \). Clearly the probability that player \( i \) chooses pure strategy \( j \) is increasing in the subconscious utility associated with \( j \) (i.e., \( U^i_j(P^i) \)) but decreasing in the subconscious utility associated with any other pure strategy. These choice probabilities reflect boundedly rational behavior, because the player generally does not choose what is best for himself in the sense of expected utility maximization. Nevertheless, the choice probabilities do exhibit a tendency toward subconscious utility maximization because they are ranked in the same order as the subconscious utilities.
In place of Eq. (2) consider

$$p_i(P^i) = \frac{f[U_i^i(P^i), \mu_i]}{\sum_{j=1}^m f[U_j^i(P^i), \mu_j]}.$$  (3)

If $f$ were continuous, bounded for each $\mu_i \in (0, \infty)$, and $f[U_i^i(P^i), \mu_i] > 0$ for all admissible values of $(U_i^i(P^i), \mu_i)$, then the choice probabilities would be well defined and also existence of equilibrium would be assured. If $f[U_i^i(P^i), \mu_i]/f[U_j^i(P^i), \mu_j] \to 0$ as $\mu_i \to \infty$ whenever $U_i^i(P^i) < U_j^i(P^i)$ then convergence to Nash equilibrium as the $\mu_i \to \infty$ would be assured. For these results see McKelvey and Palfrey (1995). Throughout most of their paper McKelvey and Palfrey (1995) use $f[U_i^i(P^i), \mu_i] = \exp(\mu U_i^i(P^i))$. Their formulation is invariant to changes in the utility origin but not to linear (i.e., scale) transformations and it will accommodate negative utilities, because $\exp(u)$ is always positive. The formulation we use is invariant to linear transformations, but not to changes in the utility origin. As noted above, all results go through for either version and for many other formulations as well.

3. BOUNDEDLY RATIONAL NASH EQUILIBRIUM

In this section we define the boundedly rational Nash equilibrium and discuss its existence and asymptotic behavior. In essence, this equilibrium is a mixed strategy profile under which the strategy of each player is a vector of discrete choice probabilities which is a random choice best reply to the choice probabilities of the remaining players. Thus these choice probabilities are calculated for each player $i$ using the probability distribution over $A_i^i$ that is induced (implied) by the strategies of the other players; therefore, equilibrium is a profile of mutually consistent choice probability distributions. The next step in this section is to introduce mixed strategies and relate them formally to the choice probabilities.

As $p^i(P^i) = (p_1^i(P^i), \ldots, p_m^i(P^i))$ characterizes the behavior of player $i$, it can be viewed as the mixed strategy he selects. The set of mixed strategies for player $i$ is

$$S^i = \left\{ s^i \in A^i_+ \left| \sum_{a^i} s^i_{a^i} = 1 \right. \right\}.$$  

The set of mixed strategy profiles is $S = \times_{i=1}^n S^i$ and the set of mixed strategy profiles for all players except $i$ is $S^{-i} = \times_{j \neq i} S^j$. A mixed strategy profile for the players $N \setminus \{i\}$, $s^{-i} \in S^{-i}$, induces a probability distribution
for each \( a^{-i} = (a^1, \ldots, a^{i-1}, a^{i+1}, \ldots, a^n) \in A^{-i}. \) A mixed strategy \( s^i \) of player \( i \) is consistent with the mixed strategy profile \( s^{-i} \) being followed by the other players when it represents the logit choice probabilities relative to the probability distribution \( \Pi'(\cdot | s^{-i}) \). Our definition of *boundedly rational Nash equilibrium* is that each player’s mixed strategy be consistent in this sense. That is

**DEFINITION 1.** A mixed strategy profile \( v = (v^1, \ldots, v^n) \in S \) is a boundedly rational Nash equilibrium if, for each \( i \in N, \) and each \( j = 1, \ldots, m, \)

\[
v_j^i = p_j^i \left( \Pi'(\cdot | v^{-i}) \right) = \frac{\left[ U_j^i \left( \Pi'(\cdot | v^{-i}) \right) \right]^{\mu^i}}{\sum_{k=1}^{m} \left[ U_k^i \left( \Pi'(\cdot | v^{-i}) \right) \right]^{\mu^i}}.
\]

In effect, we assume that player \( i \) acts as if he were playing a game against nature which is specific to him and is determined by \( s^{-i}. \) It is nonetheless possible to define *boundedly rational best reply functions*. Such a function for player \( i \) is merely the strategy defined by Eq. (2) that is obtained from the probability distribution \( \Pi'(\cdot | s^{-i}) \). That is, \( \beta^i(s) = (\beta_1^i(s), \ldots, \beta_n^i(s)) \), the boundedly rational best reply of player \( i \) to \( s \), is defined by

\[
\beta_j^i(s) = p_j^i \left( \Pi'(\cdot | s^{-i}) \right) = \frac{\left[ U_j^i \left( \Pi'(\cdot | s^{-i}) \right) \right]^{\mu^i}}{\sum_{k=1}^{m} \left[ U_k^i \left( \Pi'(\cdot | s^{-i}) \right) \right]^{\mu^i}}.
\]

and the *boundedly rational best reply function* for the game is

\[
\beta(s) = (\beta_1(s), \ldots, \beta_n(s)).
\]

Clearly \( \beta \) is continuous in \( s \) and existence of a *boundedly rational Nash equilibrium* in a finite game follows easily.

**THEOREM 1** (McKelvey and Palfrey, 1995). *A finite game \( (N, A, u) \) has a boundedly rational Nash equilibrium.*

There is a fundamental question concerning the attainment of such an equilibrium. Suppose \( n \) players are in a one shot game, as described above. Where do the choice probabilities come from? Admittedly, this is not clear; the probability of choosing a particular pure strategy depends on the subconscious utility attached to that strategy as compared with the subconscious utility attached to the other pure strategies. Meanwhile, the subcon-
ous utility of a strategy cannot be computed without knowing the mixed strategies of the other players. Hence, the equilibrium can be regarded as a set of consistency conditions; that is, a player’s choice probabilities are correct in the sense of reflecting what the player must do as a random choice player when the choice probabilities are a boundedly rational best reply to the strategies of the other players.

In contrast, a repeated game may be postulated in which the same set of players will play precisely the same game over and over again. As these players are boundedly rational, it is not unreasonable to suppose that they play myopically according to the random choice model that we have supposed. Each player might begin by using an arbitrary (completely) mixed strategy. Then, each player’s chosen strategy may adjust over time in light of the observed play and converge to a boundedly rational Nash equilibrium. This topic is pursued in Section 4.

Turning to the asymptotic behavior of equilibrium, we are concerned with the nature of equilibrium as the players become increasingly rational. This is captured by permitting the \( \mu^i \) to go to infinity. As the parameters of all players in a game go to infinity, equilibrium behavior converges to a conventional Nash equilibrium.

**Theorem 2** (McKelvey and Palfrey, 1995). Let \( \Gamma = (N, A, u) \) be a finite game and let \( \{ \mu'^i \}_{i=1}^m \) be a sequence of points in \( \mathcal{F}_+^m \) such that \( \min \mu'^i \to \infty \) as \( r \to \infty \). For each \( \mu'^i \) let \( v'^i \) be a fixed point of \( \beta(\cdot | \mu'^i) \). Then the sequence \( \{ v'^i \}_{i=1}^m \) has at least one point of accumulation, \( v^0 \), and any such point of accumulation is a Nash equilibrium of \( \Gamma \).

One might guess that \( v^0 \) would be a perfect equilibrium, because, when all elements of \( A \) are positive, then all \( v'^i \) are completely mixed. This conjecture is wrong. Another conjecture is that any Nash equilibrium can be obtained as the limit of a sequence of boundedly rational Nash equilibria. This, too, is false. McKelvey and Palfrey (1995) provide counterexamples to these conjectures.

It is interesting to see the effect of a change in the utility origin on the boundedly rational Nash equilibrium for the particular random choice formulation used here and to compare it to the Nash equilibrium. In general, as the value of a positive constant added to the subconscious utility function is increased, the choice probabilities move nearer to \( 1/m \). This is illustrated for the prisoners’ dilemma shown in Table I. Suppose that \( a \) in the table is nonnegative. The Nash equilibrium is, of course, well known to be (defect, defect). We calculate the boundedly rational Nash equilibrium when \( \mu_1 = \mu_2 = 1 \); however, qualitatively similar results would hold for other values of the \( \mu^i \). If player 2 selects defect with probability \( p \), the utility of defect for player 1 is \( u_1 = p(5 + a) + (1 - p)(15 + a) = 15 + a - 10p \) and the utility of cooperate is \( u_2 = pa + (1 - p)(10 + a) \), etc.
TABLE I
The Prisoner’s Dilemma

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Defect</th>
<th>Cooperate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defect</td>
<td>$5 + a, 5 + a$</td>
<td>$15 + a, a$</td>
</tr>
<tr>
<td>Cooperate</td>
<td>$a, 15 + a$</td>
<td>$10 + a, 10 + a$</td>
</tr>
</tbody>
</table>

Thus the boundedly rational best reply for player 1 is $w_1 = \frac{10 + a - 10p}{15 - q}$. At the equilibrium, the probability of choosing defect must satisfy

$$\omega(a) = \frac{15 + a - 10\omega(a)}{25 + 2a - 20\omega(a)}, \quad (4)$$

which can be solved for $\omega(a)$:

$$\omega(a) = \frac{35 + 2a - \sqrt{(15 + 2a)^2 - 200}}{40}. \quad (5)$$

For $a > 0$ the only relevant root of Eq. (4) is the one shown in Eq. (5), which is illustrated in Fig. 1. However, for $a = 0$ both roots are equilibria; they are 3/4 and 1. The smaller root is the limit of Eq. (5) as $a \rightarrow 0$. The relationship between the boundedly rational equilibrium and the Nash equilibrium reflects their respective natures in a very clear way. First, apart from the larger root when $a = 0$, the boundedly rational equilibrium always assigns positive probability to both pure strategies. Furthermore, when $a > 0$, the probability assigned to defect never exceeds 3/4. As $a$ increases, the probability falls with the limit being 1/2. This expresses the increasing relative similarity of the payoffs as $a$ increases; that is, as $a$ increases, the difference between the largest and smallest payoff, divided by the smallest payoff, goes to zero. This relationship is central to the determination of the choice probabilities, driving them to equality; however, it is irrelevant to the Nash equilibrium on which $a$ has no effect. For the Nash equilibrium only the differences among payoffs matter and, for any value of $a$, equilibrium remains uniquely (defect, defect).

This example illustrates that the particular choice of the function $f$ (see Eq. (3)) has an important effect on the character of boundedly rational choice. The sensitivity of choice probabilities to the relative sizes of the subconscious utilities is appealing in some circumstances, but not necessarily in all. It has, for example, been drawn to our attention that consumers
will go to some considerable trouble to purchase an item (e.g., an electric shaver) at $45 rather than pay $50, but, at the same time, they will typically go to no trouble to purchase an $18,000 item for $17,900 (e.g., an automobile).\footnote{Personal communication to Friedman from R. Selten.}

4. A DYNAMIC ADJUSTMENT PROCESS

This section deals with global and local stability properties of the boundedly rational Nash equilibrium for myopic players who engage in the same game repeatedly. It is shown below that equilibrium is globally stable for relatively small values of the $\mu_i$. For arbitrarily large values of the $\mu_i$ local stability depends upon the nature of the equilibrium. Strict Nash equilibria (i.e., equilibria at which each player's best reply is unique) are locally stable, but other equilibria are problematic. The stability results are based upon the best reply function being a contraction. For small $\mu_i$ it is a contraction globally. In an open neighborhood of a strict equilibrium, it is a contraction even as the $\mu_i \to \infty$; however, in most games the best reply function will cease being a contraction on its domain as $\mu_i \to \infty$. Indeed, in the limit it will not satisfy any global Lipschitz condition (except for games such as the prisoners' dilemma, in which there is a unique Nash equilibrium that also is strict).

\footnote{Personal communication to Friedman from R. Selten.}
4.1. Global Stability

Suppose that the game is repeated in each time period \( t \). In each play of the game, the players are myopic and boundedly rational; however, each has some sense of the behavior of other players that is a consequence of the observed choices of the other players over time. Assume that at any time \( t \geq 1 \) all players have beliefs based upon the empirical distributions \( \bar{P}_i^t = (\bar{P}_i^1, \ldots, \bar{P}_i^n) \) associated with each player \( i \). The empirical distribution is the average observed behavior of the player from the initial period to the present. Denote by \( \alpha^i_t \in S^i \) the degenerate probability distribution that represents actual choice of player \( i \) at time \( t \) and let \( \alpha_t = (\alpha^i_t, \ldots, \alpha^n_t) \in S \); thus, \( \alpha^i_t = (\alpha^i_{1,t}, \ldots, \alpha^i_{m,t}) \) is a vector with all entries zero except for an entry of one for the coordinate corresponding to the actual pure action selected in period \( t \). Player \( i \) will typically randomize at time \( t \); however, \( \alpha^i_t \) is the realization of that randomization and is observable by all players. The empirical distribution is then

\[
\bar{P}_i^t = \sum_{\tau=0}^{t-1} \frac{\alpha^i_{\tau}}{t}
\]

for \( t \geq 1 \) and is used at time \( t \) to calculate the subjective utilities prior to taking action. Let \( \bar{P}_i = (\bar{P}_i^1, \ldots, \bar{P}_i^n) \) and \( \bar{P}_i^{-1} = (\bar{P}_i^1, \ldots, \bar{P}_i^{t-1}, \bar{P}_i^{t+1}, \ldots, \bar{P}_i^n) \). The empirical distributions generate the beliefs that players have about one another because \( \bar{P}_i^{-1} \) induces a probability distribution over \( A^{-i} \). In each period, player \( i \) chooses with probabilities that constitute the boundedly rational best reply to \( \bar{P}_i^{-1} \).

This is the fictitious play proposed by Brown (1951) and Robinson (1951) except that the best reply behavior is boundedly rational instead of being Nash. As Shapley (1964) has shown, fictitious play does not generally converge for fully rational players in finite games; however, it will be seen below that it converges for boundedly rational players in a large class of games. The key difference between best replies for fully rational players and best replies for boundedly rational players is that the best reply mapping for a boundedly rational player is always a single-valued mapping and, for a given game and given \( \mu_i \), will satisfy a Lipschitz condition. In contrast, the best reply mapping for fully rational players is not always

\(^2\) As \( \bar{P}_i^t \) is the empirical distribution of player \( i \), \( \bar{P}_i \) is a profile of empirical distributions. We will abuse terminology by referring to \( \bar{P}_i \) as an empirical distribution in order to avoid cumbersome language.
single valued and will not, in general, admit a selection that is continuous (or, consequently, Lipschitz). In the proof of Theorem 3, a crucial condition is that the boundedly rational best reply function is a contraction (i.e., Lipschitz with ratio less than one). Because players observe realizations of mixed strategies rather than mixed strategies themselves, the proof of convergence involves a good deal more than this Lipschitz condition. Specifically, it must be shown that the empirical distribution converges and, in the limit, that it equals the strategy profile used by the players.

To state more precisely what is proved and how it is proved, a few additional definitions are needed. The following distance measure on $S^i$ is used: $\rho(x^i, y^i) = \sum_{k=1}^{m} |x^i_k - y^i_k|/2m$. The distance measure on $S$ is given by $\rho(x, y) = \sum_{i=1}^{n} \rho(x^i, y^i)/n$. $\bar{P}^{-i}$ induces a probability distribution over $A^{-i}$ where the probability associated with $a^{-i}$ is $\Pi_{j \neq i} \bar{P}^j_{t_i}$. Suppose a function $f$ maps $S$ to itself; that is $f: S \rightarrow S$. The function $f$ satisfies a Lipschitz condition with ratio $\lambda$ if for any $s, s' \in S$ we have $\rho(f(s), f(s')) \leq \lambda \rho(s, s')$. When $f$ satisfies a Lipschitz condition with ratio $\lambda < 1$ it is a contraction.

We are interested in the following questions: (i) Will $\bar{P}^i, i \in N$, converge in probability as $t \rightarrow \infty$? If the answer to the preceding question is yes, then clearly the strategy $s^i_t$ of player $i$ at time $t$, given by

$$s^i_t = \beta^i_t(\bar{P}) = p^i_t(\Pi^i(\cdot | \bar{P}^{-i})) = \frac{\left[ U^i_t(\Pi^i(\cdot | \bar{P}^{-i})) \right]^{\mu^i}}{\sum_{k=1}^{m} \left[ U^i_t(\Pi^i(\cdot | \bar{P}^{-i})) \right]^{\mu^i}},$$

must converge to some $\hat{s}^i$. However, (ii) Is the resulting $\hat{s} = (\hat{s}^1, \ldots, \hat{s}^n) \in S$ a fixed point of the equation system (7) and consequently a boundedly rational Nash equilibrium?

We address these questions in turn. In a series of lemmas it is shown below that if the best reply function $\beta$ is a contraction, then the boundedly rational Nash equilibrium is unique and fictitious play converges in probability to the equilibrium strategy profile. Let $\bar{u} = \max_{a \in A} u^i_a$ and $\underline{u} = \min_{a \in A} u^i_a$.

**Lemma 1.** For all $i \in N$, all $j = 1, \ldots, m$, $\bar{u}, \underline{u} > 0$, and finite, nonnegative $\mu^i$, $\beta^i$ satisfies a Lipschitz condition.

**Proof.** It is evident that each $U^i_j$ is Lipschitz; therefore, for positive $\mu^i$, $[U^i_j]^{\mu^i}$ is also Lipschitz. Denote by $\theta$ the largest of the $m$ Lipschitz constants for the $[U^i_j]^{\mu^i}$, let $\bar{P}$ and $\bar{P}'$ be two arbitrary empirical distributions, and let $q = \Pi(\cdot | \bar{P}^{-i})$ and $q' = \Pi(\cdot | \bar{P}'^{-i})$. Then $\| [U^i_j(q)]^{\mu^i} - \| [U^i_j(q')]}^{\mu^i}$.
which is clearly finite; therefore, each $\beta_i^j$ is Lipschitz with ratio not greater than $2m^\mu/(mu^\mu)^2$. 

Proving convergence of the empirical distribution requires results that are developed next in Lemmas 2 and 3.

**Lemma 2.** For arbitrary $P$, and arbitrary $\delta > 0$ let $P_{t+\nu}$ be the empirical distribution that holds at time $t + \nu$ when the empirical distribution at time $t$ is $P$, and choices at each time from $t$ to $t + \nu$ are governed by $\beta$. If $\nu/(t + \nu) \leq \delta$ then $\rho(P_t, P_{t+\nu}) < \delta$.

**Proof.** Denote by $\alpha_\tau$ the realization of the players' choices at times $\tau = t, \ldots, t + \nu - 1$. Then

$$P_{t+\nu} = \frac{t}{t + \nu} P_t + \frac{\nu}{t + \nu} \sum_{\tau=0}^{\nu-1} \alpha_{t+\tau}.$$ 

By choosing $\nu$ and $t$ so that $\nu/(t + \nu) = \delta$ it is clear the lemma holds. 

To prove convergence in probability of the empirical distribution it is necessary to establish some results that draw upon the Bernoulli Law of Large Numbers and the Covering Rule (see Loève 1963, p. 14 and 16). Let $p = (p_1, \ldots, p_m)$ be a probability distribution over $m$ outcomes and let $\pi_k = (\pi_{1k}, \ldots, \pi_{mk})$ be the empirical distribution resulting from $k$ independent observations. We may treat the occurrence of one particular element of $\{1, \ldots, m\}$, say element $j$, as a Bernoulli process. Then the Bernoulli Law of Large Numbers states that, for any $\epsilon > 0$ and any
positive integer $k$,  
\[
\Pr[|\pi_{jk} - p_j| < \varepsilon] > 1 - \frac{p_j(1 - p_j)}{k\varepsilon^2}
\]  
(8)

or  
\[
\Pr[|\pi_{jk} - p_j| > \varepsilon] < \frac{p_j(1 - p_j)}{k\varepsilon^2}.
\]  
(9)

Now consider the probability that $|\pi_{jk} - p_j|\leq \varepsilon$ for all $j$. This can be addressed with the Covering Rule in the following way. The probability that one particular $|\pi_{jk} - p_j| > \varepsilon$ is no greater than $p_j(1 - p_j)/k\varepsilon^2$ and the Covering Rule states that the probability that Eq. (9) holds for at least one $j \in \{1, \ldots, m\}$ is no more than $\sum_{j=1}^{m} p_j(1 - p_j)/k\varepsilon^2$. The latter is bounded above by $1/k\varepsilon^2$. Consequently,  
\[
\Pr\left[\max_{1 \leq j \leq m} |\pi_{jk} - p_j| < \varepsilon\right] > 1 - \frac{1}{k\varepsilon^2}.
\]

It is now possible to consider an empirical distribution $\pi_k$ that is the result of observations drawn from a sequence of probability distributions that are all close to one another.

**Lemma 3.** Let $(p_k)_{k=1}^{\infty}$ be a sequence of discrete probability distributions on $\{1, \ldots, m\}$ satisfying the condition that $p_{jk} - \delta \leq p_{jk} \leq p_{jk} + \delta$ for $j \in \{1, \ldots, m\}$ and all positive integers $k$. Suppose that one observation is drawn from each member of the sequence and that $\pi_k$ is the empirical distribution that results from the observations drawn, one each, from $p_1, \ldots, p_k$. Then for any $\varepsilon > 0$ and any positive integer $k$,  
\[
\Pr\left[\max_{1 \leq j \leq m} |\pi_{kj} - p_{1j}| < \delta + \varepsilon\right] > 1 - 1/k\varepsilon^2.
\]  
(10)

**Proof.** If the empirical distribution resulted from observations drawn only from $p_1$ then clearly for a particular $j$, $p_{1j} - \varepsilon < \pi_j < p_{1j} + \varepsilon$ with probability $1 - p_{1j}(1 - p_{1j})/k\varepsilon^2$. However, when the empirical distribution is drawn with varying probabilities that are never less than $p_{jk} - \delta$ and never more than $p_{jk} + \delta$, the probability of $p_{1j} - \varepsilon - \delta < \pi_j < p_{1j} + \varepsilon + \delta$ must be at least $1 - p_{1j}(1 - p_{1j})/k\varepsilon^2$ because the sample, drawn in this latter way, has no higher probability of being below the lower bound and no higher probability of being above the upper bound, as compared with the standard case of Bernoulli trials. Therefore, Eq. (10) follows by application of the Covering Rule.  


Convergence of the empirical distribution is proved next in Theorem 3. The proof relies on $\beta$ being a contraction. Recall that $\beta$ transforms the empirical distribution at time $t$ into the choice probabilities of time $t$, and not into the empirical distribution of time $t + 1$. Furthermore, these choice probabilities are not directly observed; players see realizations. Therefore, we look at the relationship between $\bar{P}_t$ and $\bar{P}_{t+\nu}$ with the times $t$ and $t + \nu$ carefully chosen so that the choice probabilities hardly change (i.e., the $s_j$ for $t = \ldots, t + \nu$ hardly change). This is a way to compensate for the inability of any player $i$ to observe any $s_j, i \neq j$, directly and permits a proof of convergence that then proceeds in much the same way that the proof would be carried out if the $s_j$ were directly observable.

**Theorem 3.** If $\beta$ satisfies a Lipschitz condition with ratio $\theta < 1$ then the empirical distribution converges in probability.

**Proof.** Suppose that $\bar{P}_t$ and $\bar{P}_t'$ are empirical distributions and denote by $\bar{P}_{t+\nu}$ and $\bar{P}_{t+\nu}'$, respectively, the empirical distributions to which $\bar{P}_t$ and $\bar{P}_t'$ lead; $\delta = \nu/(t + \nu)$. Let $s_{\tau} = \beta(P_{\tau})$ and $s_{\tau}' = \beta(P_{\tau}')$ for $\tau = t, \ldots, t + \nu - 1$, where $\bar{P}_t$ and $\bar{P}_t'$ have the obvious meanings. Finally, let $\pi_{t, \nu} = \sum_{s_{t+\nu}} P_{t+\nu}(s_{t+\nu})/\nu$ and $\pi_{t, \nu}' = \sum_{s_{t+\nu}'} P_{t+\nu}'(s_{t+\nu}')/\nu$ so that $\bar{P}_{t+\nu} = (1 - \delta)\bar{P}_t + \delta P_{t+\nu}$ and $\bar{P}_{t+\nu}' = (1 - \delta)\bar{P}_t' + \delta P_{t+\nu}'$. Then

$$\rho(\bar{P}_{t+\nu}, \bar{P}_{t+\nu}') \leq (1 - \delta)\rho(\bar{P}_t, \bar{P}_t') + \delta \rho(\pi_{t, \nu}, \pi_{t, \nu}')$$

with probability no less than $1 - 1/\nu^2$ and $\rho(\pi_{t, \nu}', \pi_{t, \nu}) < \delta + \varepsilon$ with probability of at least $1 - 1/\nu^2$ so that the Covering Rule implies that $\rho(\pi_{t, \nu}', \pi_{t, \nu}) < \delta + \varepsilon$ with probability of at least $1 - 2/\nu^2$. Using the Lipschitz condition $\theta$ on $\beta$, this, in turn, implies

$$\rho(\bar{P}_{t+\nu}, \bar{P}_{t+\nu}') \leq (1 - \delta)\rho(\bar{P}_t, \bar{P}_t') + \delta \rho(\pi_{t, \nu}, \pi_{t, \nu}') + 2(\delta + \varepsilon)$$

with probability of at least $1 - 2/\nu^2$. Thus $\rho(\bar{P}_{t+\nu}, \bar{P}_{t+\nu} < \rho(\bar{P}_t, \bar{P}_t')$ with probability of at least $1 - 2/\nu^2$ if the right-hand side of Eq. (11) is less than $\rho(\bar{P}_t, \bar{P}_t')$. The latter is equivalent to

$$\frac{2(\delta + \varepsilon)}{1 - \theta} < \rho(\bar{P}_t, \bar{P}_t')$$

What Eq. (12) says is that when two empirical distributions are farther apart than $2(\delta + \varepsilon)/(1 - \theta)$, then, with probability $1 - 2/\nu^2$ their successor distributions at time $t + \nu$ will be closer together. The remainder of the proof will show two things. First, it is possible to choose a succession of time periods, $t, t + \nu_1, t + \nu_1 + \nu_2, \ldots$ such that, with probability bounded away from zero, the empirical distributions stemming from $\bar{P}_t$ get progressively closer to those stemming from $\bar{P}_0$ as long as Eq. (12) holds. This
really means that eventually, with positive probability, empirical distributions converge to a neighborhood whose size is governed by Eq. (12). The second part of the remainder of the proof is to show that the left-hand side of Eq. (12) can be made arbitrarily small and that the probability of converging to an arbitrarily small neighborhood is almost one. In the limit, the neighborhood collapses to a point and the probability goes to one.

By the Covering Rule, Eq. (11) holds simultaneously for all players with probability 1 \(\frac{2}{n}\mu^2\). Define \(v_t = v\) and \(t_1 = t\) and then define both \(v_k\) and \(t_k\) recursively by \(v_k = v(t + v/t)^{k-1}\) and \(r_k = t_k - t_{k-1} + v_{k-1}\). It is easily verified that \(v_k/(t_k + v_k) = v/(t + v)\) \(= \delta\) for all \(k\). Consider now a sequence \((P_i)_{t=0}^\infty\) of empirical distributions emanating from some initial \(P_0\) and a second sequence \((P'_i)_{t=0}^\infty\) emanating from \(P_0\). From Eq. (11),

\[
\rho(P_i, P'_i) \leq (1 - \delta) \rho(P_i, P'_i) + \delta \left( \theta \rho(P_i, P'_i) + 2(\delta + \epsilon) \right)
\]

(13)

for all \(i \in N\) with probability of at least \(1 - 2n/\nu_k \epsilon^2\). Consequently, by the Covering Rule, Eq. (13) holds for all \(k\) with probability of at least

\[
1 - \sum_{k=1}^{\infty} 2n/\nu_k \epsilon^2 = 1 - 2n/\left[(1 - \delta) \nu v^2\right] = 1 - 2n/\delta \epsilon^2.
\]

Thus with probability no smaller than \(1 - 2n/\delta \epsilon^2\) any pair of empirical distributions will eventually satisfy the condition that \(\rho(P_i, P'_i) \leq 2(\delta + \epsilon)/(1 - \theta)\). Clearly, both \(\delta\) and \(\epsilon\) can be chosen arbitrarily small and, given \(\delta\) and \(\epsilon\), \(t\) can be selected to make \(2n/\delta \epsilon^2\) arbitrarily small. Thus, in the limit, Eq. (12) will hold for all \(i \in N\) and \(t = t_1, t_2, \ldots\) with probability one so that \(P_i\) converges in probability to the fixed point of \(\beta\) if \(\beta\) satisfies a Lipschitz condition with ratio \(\theta < 1\).

It remains to show that fictitious play converges in a large class of games. This is done in Theorem 4 by showing that the derivatives of each \(\beta^i\) are continuous in \(\mu^i\) and go to zero as \(\mu^i\) goes to zero. This, in turn, implies that \(\beta^i\) is a contraction for some measurable interval of \(\mu^i\) values whose lower limit is zero and upper limit is positive.

**Theorem 4.** For any finite game for which \(u^i_j > 0\) for all \(i \in N, j = 1, \ldots, m,\) and \(a \in A\), there is a set \(M = \{ \mu \in \mathbb{R}^m \mid 0 \leq \mu^i < M, i \in N\}\) such that if \(\mu \in M\) then \(\beta\) is a contraction. Therefore, if \(\mu \in M\), fictitious play converges to a boundedly rational Nash equilibrium which is unique.

**Proof.** If \(\mu = 0\) then players always choose each pure strategy with probability \(1/m\) so that \(\beta\) is trivially a contraction. It is clear that \(\beta\) is
continuous in \( \mu \); indeed, it is continuously differentiable in \( \mu \):

\[
\frac{\partial \beta_i^j(\bar{P})}{\partial \bar{P}_i^j} = \mu^{-1} \beta_i^j \left[ \left( \frac{\partial U_i^j}{\partial \bar{P}_i^j} \right) \left( \sum_k \beta_k U_i^k \right)^{-1} - \left( \sum_k \beta_k \right)^{-1} \right]
\]

The absolute value of \( \frac{\partial \beta_i^j(\bar{P})}{\partial \bar{P}_i^j} \) is bounded above by \( \mu (\bar{n}^2 - \bar{u}^2) / \bar{u} \) which goes to zero as \( \mu \to 0 \) and therefore implies the existence of a nonempty set \( M \) on which \( \beta \) is a contraction. When \( \beta \) is a contraction, the limit of the empirical distribution is the unique fixed point of \( \beta \); hence it is a boundedly rational Nash equilibrium and is unique.

Shapley’s (1964) example in which fictitious play fails to converge makes use of special properties of the Nash best reply mapping. That mapping is, in general, set valued and does not admit a continuous selection. This can be seen by recalling the conditions under which a mixed strategy is a Nash best reply for a player. The pure strategies in the support of the mixed strategy must, themselves, all be best replies. The slightest alteration in the other players’ strategies, will, in general, cause one member of the support to remain the unique best reply. Which member that is will depend upon the exact nature of the slight alteration. For fictitious play this means that a tiny change in the empirical distribution will sometimes cause a very large change in the best reply of a player. For example, in the game of matching pennies, suppose a player wins when the coins match and let the opposing player’s probability of heads be \( q \). Clearly the player’s unique best reply is heads when \( q > 0.5 \) and tails when \( q < 0.5 \). If \( q \) is close to 0.5 then a slight change in the empirical distribution could cause the best reply to switch from heads to tails. Shapley exploits this characteristic of best replies to construct his insightful and instructive example. It is possible to approximate the best reply correspondence in the matching pennies game with a single-valued function that satisfies a Lipschitz condition; however, as the approximation converges to the true best reply mapping, the Lipschitz constant goes to infinity.

In contrast, the boundedly rational best reply mapping is a single-valued function and is also Lipschitz. That fictitious play converges for small
values of $\mu$ and cannot be shown to converge for large values is natural. First, as discussed above, convergence does not generally occur when players are fully rational, so that we should expect the same to hold for large values of $\mu$ as well. Second, in other probabilistic choice models new results regarding existence and stability of equilibria hold only when $\mu$ is not too large (see de Palma et al., 1985; Myiao and Shapiro, 1981). Whether the largest values of $\mu$ for which convergence holds may be called very rational or not very rational may depend on the particular game as well as on the point at which a given person wishes to draw the line between very and not very rational.

Indeed, it is instructive to examine convergence as a function of $\mu$ for the battle of the sexes shown in Table II. This representation of the game has all payoffs raised by one unit from conventional values; thus each player receives one when the two pick differently and when they pick the same thing, the payoffs are two and three. Let $p$ denote the probability with which woman chooses ballet and $q$ the probability with which man chooses ballet. The best reply function for man is

$$q = \frac{(1 + 2p)^\mu}{(1 + 2p)^\mu + (2 - p)^\mu}$$

and the derivative of the best reply function with respect to $p$ is

$$\frac{\partial q}{\partial p} = \frac{5\mu(1 + 2p)^{\mu-1}(2 - p)^{\mu-1}}{[(1 + 2p)^\mu + (2 - p)^\mu]^2}.$$ 

The derivative is clearly positive and is less than one for all $p \in [0, 1]$ if $\mu < 2.15$. McKelvey and Palfrey show that, as $\mu$ goes from zero to infinity,
a unique path of equilibria is traced. In the present game this unique path will end at the mixed strategy equilibrium; therefore it is not surprising that the Lipschitz condition that is satisfied by the best reply function becomes larger as \( m \) increases.

### 4.2. Local Stability

Two results relating to stability are proved in Theorem 5 below. The first is that a strict Nash equilibrium has a nonempty open neighborhood in which boundedly rational Nash equilibria are stable for all \( \mu' \) above some critical finite value. The second is that, for any game in which the Nash best reply fails to be a single-valued mapping everywhere on its domain, the boundedly rational best reply mapping cannot be globally Lipschitz; that is, if one player \( i \) has two Nash best replies at merely one \( s^{-i} \) then the boundedly rational best reply function is not Lipschitz for large \( \mu' \).

Suppose a game \((N, A, u)\) has the property that, for at least one player \( i \), there is at least one \( s^{-i} \in S^{-i} \) with respect to which player \( i \) has more than one but fewer than \( m \) best replies. Denote the set of finite games having this property by \( \mathcal{F} \). It is shown in Theorem 5 that the best reply function cannot be a contraction on the entire domain \( S \) in any game in \( \mathcal{F} \). It is also shown that a strict Nash equilibrium of any finite game has an open neighborhood in which the best reply function is a contraction. Consequently, boundedly rational Nash equilibrium is stable in such a neighborhood for large \( \mu' \). Let \( A^{-ik} = \times_{i \in N \setminus \{i, k\}} A_i \).

**Theorem 5.** Let \((N, A, u)\) be a finite game. (i) If \( s^* \) is a strict Nash equilibrium of \((N, A, u)\) then there is a finite \( \mu^x \) and an \( \varepsilon \) neighborhood of \( s^* \) such that the boundedly rational best reply function is a contraction in this neighborhood and any boundedly rational Nash equilibrium is stable in this neighborhood for all \( \mu \) such that \( \mu' > \mu^x \), \( i \in N \). (ii) For any \((N, A, u) \in \mathcal{F}\) the Lipschitz constant \( \lambda \) satisfied by the boundedly rational best reply function goes to infinity as \( \mu \to \infty \).

**Proof.** Recall that \( \sum_{j=1}^m \beta_j^i = 1 \) and use \( \beta_m^i = 1 - \sum_{j=1}^{m-1} \beta_j^i \). Then the derivatives of the best reply function are

\[
\frac{\partial \beta_j^i}{\partial s_k^{u^i}} = \mu^i \left( \sum_{l=1}^m \beta_l^i \right)^{\mu' - 1} \frac{\partial U_j^i}{\partial s_k^{u^i}} - \left( \sum_{l=1}^m \beta_l^i \right)^{\mu'} \left( \sum_{l=1}^m \beta_l^i \right)^{\mu' - 1} \frac{\partial U_j^i}{\partial s_k^{u^i}},
\]

(14)
where

\[
\frac{\partial U^i_j}{\partial s_{a_k}^j} = \sum_{a^{-ik} \in A^{-ik}} u_{j, a^k, a^{-ik}} \prod_{l \in N \backslash \{u, k\}} s_{a^{-ik}}^l. \tag{15}
\]

Suppose \( s^* \) is a strict Nash equilibrium, the equilibrium pure strategy of player \( i \) is \( j \) (without loss of generality, \( j \neq m \)) and evaluated at \( s^* \), \( U^i_j > U^i_k + 2\delta \) for some \( \delta > 0 \) and all \( k \neq j \). Denote by \( S_\delta \) an open neighborhood of \( s^* \) in which \( U^i_j > U^i_k + \delta \) for all \( k \neq j \). Clearly such a neighborhood exists. Choose \( s \) in this neighborhood and then note that the absolute value of \( \frac{\partial \beta^i_j}{\partial s_{a_k}^j} \) is bounded above by

\[
\frac{\mu^i \bar{\mu} [U^i_j (U^i_j - \delta)]^{\mu^i-1} (m - 1) (2U^i_j - \delta)}{(U^i_j)^{2\mu^i}} \tag{16}
\]

where \( \bar{\mu} = \max_{a \in A} u_{a}^i \). It is clear that Eq. (16) goes to zero as \( \mu^i \to \infty \), which establishes that, for large \( \mu \), the best reply function is a contraction in a neighborhood of \( s^* \). It is also obvious that, for sufficiently large \( \mu \), the best reply function must map a nonempty, open subset of \( S_\delta \) into itself. This establishes (i).

To show (ii) it suffices to prove that the absolute value of Eq. (14) goes to infinity as \( \mu^i \to \infty \). Choose some \( s \in S \) such that both pure strategies \( j \) and \( r \) are best replies, but strategy \( m \) is not a best reply for player \( i \). By definition of \( \mathcal{G} \), for any game \( (N, A, u) \in \mathcal{G} \) there must be a player \( i \) and strategy profile \( s \) that satisfy these conditions.\(^3\) Equation (14) can be rewritten

\[
\frac{\partial \beta^i_j}{\partial s_{a_k}^j} = \mu^i \left[ (U^i_m)^{\mu^i} - (U^i_j)^{\mu^i} \right] \sum_{l \neq i, m} (U^i_l)^{\mu^i-1} \frac{\partial U^i_j}{\partial s_{a_k}^j} + \left( U^i_j U^i_m \right)^{\mu^i-1} \left( U^i_m \frac{\partial U^i_j}{\partial s_{a_k}^j} - U^i_j \frac{\partial U^i_m}{\partial s_{a_k}^j} \right) \left[ \sum_{l=1}^{m} (U^i_l)^{\mu^i} \right]^2, \tag{17}
\]

\(^3\) Of course, if \( m \) happens to be a best reply, the strategies may be renumbered so that \( m \) is no longer a best reply.
As $\mu^i$ grows, the absolute value of Eq. (17) is greater than

$$\frac{\mu^i (U_i^j)^{\mu^i} \frac{\partial U_i^j}{\partial s_{a_k}^i}}{\left[ \sum_{j=1}^m (U_i^j)^{\mu^i} \right]^2} = \frac{\mu^i (U_i^j)^{2\mu^i-1} \frac{\partial U_i^j}{\partial s_{a_k}^i}}{\left[ \sum_{j=1}^m (U_i^j)^{\mu^i} \right]^2} \geq \frac{\mu^i \frac{\partial U_i^j}{\partial s_{a_k}^i}}{n^2 (U_i^j)^{2\mu^i}} = \mu^i \frac{\partial U_i^j}{n^2 U_i^j},$$

which clearly goes to infinity as $\mu^i \to \infty$. ■

5. CONCLUSION

In this paper a method of modeling boundedly rational interactive decision has been proposed under which players are not utility maximizers. Instead they follow the random choice interpretation of discrete choice theory under which each player selects a mixed strategy in a way that shows the influence of a subconscious utility function. Although we illustrate using the multinomial logit model, the results in Sections 3 and 4 can easily be extended to a general class of discrete choice models such as those considered by McFadden (1981) where choice probabilities are (weakly) increasing in own utility and (weakly) decreasing in the other alternatives’ utilities. A great advantage of using the multinomial logit is that a single parameter in the model governs the extent to which a player’s behavior deviates from being conventional maximizing behavior. At one extreme, there is no connection between utility and choice and, at the other extreme, the connection is complete. Different players can exhibit different degrees of rationality and, for any configuration of parameter values, an analog to the Nash equilibrium exists.

Four results can be proved for the model. These are (i) existence of the boundedly rational Nash equilibrium, (ii) convergence of the equilibrium to Nash equilibrium as the rationality parameters go to infinity, (iii) sufficient conditions for fictitious play to converge to the boundedly rational Nash equilibrium, and (iv) a somewhat similar idea, expressed in a different model, has been developed in Rosen-

thal (1989).
rational Nash equilibrium, and (iv) existence of a large class of games in which the sufficient conditions for convergence will hold. Results (i) and (ii) were proved by McKelvey and Palfrey (1995), who examine single shot games having the same mathematical structure as the game we analyze, but which are based on a very different interpretation of the basis of decision making.

We believe our approach and results can help game theory to be more descriptive and applicable in practical circumstances. In this respect McKelvey and Palfrey (1995) examined empirical evidence from experiments in which there was a good deal of deviation in choice with respect to conventional best replies and found that their model did a better job of explaining the observations than the Nash equilibrium. Also, the sensitivity of the equilibrium to payoff differences appears to us as an important element of realism that differs from the Nash equilibrium. That sensitivity is not inherent in our approach to boundedly rational behavior; it is a characteristic of the particular form used here in which the choice probability show sensitivity to the origins of the utility scales.

Based on the multinomial logit model, the boundedly rational Nash equilibrium has advantages both of specificity and of flexibility. Capturing boundedly rational behavior in a way that is clearly spelled out and not ad hoc is a difficult task; we believe we have done this by using the basis borrowed from the work in discrete choice theory of Thurstone, Luce, McFadden, and others. Thus the equilibrium concept avoids arbitrary rules of thumb and is based directly on a combination of clear underlying preferences coupled with only a subconscious perception of these preferences on the part of the individual player. Flexibility is inherent in the models we have analyzed through the rationality parameter \( \mu \) which can represent everything from the extreme of purely random behavior with no connection to subconscious utility on the one hand to conventional full rationality on the other.

Another major advantage of using discrete choice theory to model bounded rationality is that it permits the use of econometric techniques for testing the actual behavior of players in games thus making it easier to evaluate the practical relevance of solutions proposed in various interactive decision making contexts.

We would be the last to say that the present work is definitive or the only fruitful way to conceive boundedly rational choice. This is a young area of investigation. Whether our particular line of analysis is very fruitful is something that only time will tell and, even if it is very fruitful, it is possible that it fits some situations well and others poorly. Other ways of modeling bounded rationality may also prove insightful. We hope that the present effort proves to be a helpful step on an interesting and important path of research.
REFERENCES


