ECONOMIC RESEARCH REPORTS

Learning, Quantal Response Equilibrium and Equilibrium in Beliefs

by

Yaw Nyarko

RR#: 2001-03 **June 2001**

C.V. Starr Center for Applied Economics

Department of Economics

Faculty of Arts and Science

New York University

269 Mercer Street, 3rd Floor

New York, New York 10003-6687

Learning, Quantal Response Equilibrium and Equilibrium in Beliefs

Yaw Nyarko*

May 1, 1998 revised June 18, 2001

Abstract

This paper makes two points. First, the modeling used in the rational (or Bayesian) learning literature can be generalized to handle the repeated shocks to preferences inherent and implicit in models of quantal response equilibria (QRE). In particular, we note that the Bayesian model and the QRE model are really not as different as often portrayed in the literature. Second, Bayesian learning under appropriate conditions therefore leads to a QRE.

Journal of Economic Literature Classification Numbers: C70, C73, D81, D82, D83, D84. Keywords: Quantal Response Equilibrium, Bayesian Learning, Game Theory

1 Introduction

This paper brings together two literatures. The first is Bayesian learning in repeated games. This literature is often referred to as rational learning. Papers in this literature include those of Jordan (1995), Kalai and Lehrer (1993), Nyarko (1998), and many others. The other literature is that on quantal response equilibria (henceforth QRE). This literature includes the those of Fudenberg and Kreps (1993) and McKelvey and Palfrey (1995) (who coined the term QRE), and many others. The QRE models are often referred to as incorporating bounded rationality or behavioral economics.

This paper makes two points. First, the modeling used in the rational (or Bayesian) learning literature can be generalized to handle the repeated shocks to preferences inherent and implicit in models of QRE. In particular, we note that the Bayesian model and the quantal response model are really not as different as often

*I thank Matthew Jackson, Richard McKelvey, Tom Palfrey and seminar particants at Cal. Tech. for helpful comments. This paper is prepared under NSF grants #SBR-98118621 and #SBR-9810723. I also thank the C.V. Starr Center at New York University for financial assistance.

portrayed in the literature. Formally, the quantal response model is typically structured to have shocks which enter the utility function, usually evolving over time in an independent and identically distributed manner. What we do here is to re-write the shocks to the utility function as Bayesian types which can be interpreted as being observed at the beginning of the infinite horizon Bayesian game.

Our second point has to do with learning. Given that we are able to embed the QRE model into a Bayesian model, we show that techniques similar to those used in the rational learning literature imply convergence to a QRE. Bayesian learning, under appropriate conditions, therefore leads to a QRE. Fudenberg and Kreps (1993) obtain the convergence to a QRE in a model similar to one studied here (what we will later call the Additive Shocks Model), but where it is assumed that learning takes place through some kind of fictitious play (i.e., "smoothed"). This paper is the Bayesian equivalent of that paper, where learning takes place via Bayesian learning. Jackson and Kalai (1997) have also studied a model similar to the one presented here with repeated shocks to preferences (which are due to different generations of individuals entering each period) and obtain related results linking the limiting play of their "recurring games" to equilibria of an underlying Bayesian game.

In the earlier literature on rational learning models, there has been quite a bit of discussion as to whether the appropriate notion of learning and convergence should be with respect to the "true" play on the one hand, or the beliefs or ex ante distribution of play on the other (see Jordan (1995) and Nyarko (1998)). In models of QRE, the true play is typically a pure action (which is contingent upon the realized shock to the utility function); the beliefs about play are a distribution. The QRE is a distribution, so is therefore in the space of beliefs, not true realized actions. In particular, it is the beliefs or ex ante play, and not the true or ex post play, which constitute a QRE. When we embed the quantal response structure into our Bayesian game setup, we therefore conclude that the appropriate notion of learning is with respect to the ex ante or beliefs; it is the beliefs which converge to a QRE.

It should be remarked in passing that Jordan (1997) and Nyarko (1994) provide results regarding the merging of ex ante play (or beliefs about play) and the empirical distribution (i.e., average play) of the true realized sample path. Applying those results here enables us to obtain results relating the limit points of empirical distributions to QRE.

2 The Model with Shocks to Payoffs

2.1 The Static Game, Γ

We begin with the static model, on which the concept of a quantal response equilibrium is defined. There are a finite number, I, of players indexed by i=1,...,I, who are playing a normal form game. Player i's action space is the finite set A^i. Define $A \equiv \prod_{i \in I} A^i$ and $A^{-i} = \prod_{j \neq i} A^j$. Given any metric space X we let $\wp(X)$ denote the set of all (Borel) probability measures on X. If for each i in a finite or countable index set C, X^i is a metric space and $q^i \in \wp(X^i)$, we define $\otimes_{i \in I} q^i$ to be the induced product measure over the product space $\prod_{i \in I} X^i$. We define $\mathcal{A}^i = \wp(A^i)$, $\mathcal{A}^{-i} \equiv \{q^{-i} = \otimes_{j \neq i} q^j : q^j \in \mathcal{A}^j \ \forall j)\}$ and $\mathcal{A} \equiv \{q = \otimes_{i \in I} q^j : q^j \in \mathcal{A}^j \ \forall j)\}$. At the beginning of the period player i observes a shock e^i to her utility function. The shock takes values in the compact set $\mathcal{E}^i$and has ex ante distribution r^i. The shocks of the different players are chosen independently of each other, so that $r = \otimes_{i \in I} r^i$ is the ex ante distribution of the vector of shocks, $\{e^i\}_{i \in I}$ on $\mathcal{E} \equiv \prod_{i \in I} \mathcal{E}^i$. The function $u^i : A \times \mathcal{E}^i \to \Re$ is player i's post-shock utility function given shock vector e^i and action profile $a = (a^i, a^{-i})$. We suppose that u^i is Borel measurable and bounded, with upper bound ū, and we normalize u^i so that it is non-negative. $u^i : A \times \mathcal{E}^i \to \Re$ is extended to $u^i : \mathcal{A} \times \mathcal{E}^i \to \Re$ via the obvious process of taking expectations. The tuple $\Gamma =< I, A, \mathcal{E}, r, \{u^i\}_{i \in I} >$ shall be referred to as the *normal form game with shocks.*

A special case of this is the *Additive Shocks Model*, where (i) the shock is a vector of the form $e^i = (e^i_n(a^i))_{a^i \in A^i}$, so that $\mathcal{E}^i = \Re^{A^i}$ (in particular, there is a scalar shock associated with each action of player i); (ii) r^i admits a continuous density function on $\mathcal{E}^i$; and (iii) the shocks enter the utility function additively: i.e., there is a *"core"* function for each player i, $w^i : A \to \Re$, such that

$$u^i(a^i, a^{-i}, e^i) = w^i(a^i, a^{-i}) + e^i(a^i) \tag{1}$$

We use the notation $\Gamma_{add} =< I, A, \{w^i\}_{i \in I}, \mathcal{E}, r >$ to refer to the Additive Shocks Model.

Define an action rule for player i to be any Borel measurable mapping $\alpha^i : \mathcal{A}^{-i} \times \mathcal{E}^i \to \mathcal{A}^i$. In particular, $\alpha^i(\pi^{-i}, e^i)$ prescribes an action or mixture in A^i when player i's shock is e^i and her belief about her opponents is π^{-i}. The action rule α^i_ε is said to be $\varepsilon -$ *optimal* against $\pi^{-i} \in \mathcal{A}^{-i}$ if $\forall a^i \in A^i$ and $r^i - a.e.\ e^i$,

$$\int_{A^{-i}} u^i(\alpha^i(\pi^{-i}, e^i), a^{-i}, e^i) d\pi^{-i} \geq \int_{A^{-i}} u^i(a^i, a^{-i}, e^i) d\pi^{-i} - \varepsilon\ . \tag{2}$$

The action rule α^i is $\varepsilon -$*optimal* if the above is true at each $\pi^{-i} \in \mathcal{A}^{-i}$. A *best response rule* is any 0-optimal action rule. We will use the notation α^i_* (resp. α^i_ε) to denote a best response rule (resp. $\varepsilon -$ *optimal* rule) for player i, and we let $\alpha_* \equiv \{\alpha^i_*\}_{i \in I}$ and $\alpha_\varepsilon \equiv \{\alpha^i_\varepsilon\}_{i \in I}$. Define $\Delta^i(\alpha^i(\pi^{-i}, .))$ to be the distribution induced over A^i by the action rule α^i at beliefs π^{-i} ; in particular, the probability assigned to action $\{a^i\}$ is

$$\Delta^i(\alpha^i(\pi^{-i}, .))(a^i) \equiv \int_{\mathcal{E}^i} \alpha^i(\pi^{-i}, e^i) \left(a^i\right) dr^i. \tag{3}$$

Fix any collection of action rules $\alpha = \{\alpha^i\}_{i \in I}$ and any $\pi = (\pi^1, ..., \pi^I) \in \mathcal{A}$. Define

$$\Delta(\alpha, \pi) \equiv \otimes_{i \in I} \Delta^i(\alpha^i(\pi^{-i}, .)). \tag{4}$$

Then $\Delta(\alpha, \pi)$ is the distribution induced over A by the profile of action rules $\alpha = \{\alpha^i\}_{i \in I}$ and beliefs of players about their opponents $\{\pi^{-i}\}_{i \in I}$.

2.2 Equilibrium Definitions

Definition 2.1. *The distribution $\pi \in \mathcal{A}$ is a quantal response equilibrium (QRE), for the normal form game with shocks $\Gamma = < I, A, E, r, \{u^i\}_{i \in I} >$ if for some best-response rules α_*, $\pi = \Delta(\alpha_*, \pi)$.*

Fudenberg and Kreps use the language "Nash equilibrium" for the above. We prefer to use different language, due to McKelvey and Palfrey. There will be many different kinds of equilibria in this paper, so we prefer dedicated language.

We now define some notions of $\varepsilon -$ *equilibria.* The quantal response equilibrium concept as defined above requires players to (i) have the correct ex ante beliefs and (ii) to best respond to those beliefs. We first define an $\varepsilon - QRE$ to be one where both of the two requirements are relaxed. We then specialize and present two other notions of $\varepsilon -$ *equilibria* which relax one of the requirements but not the other. The first, which we refer to as an $\varepsilon -$ *optimal* QRE, requires players to have correct beliefs but relaxes the requirement of best-response behavior to $\varepsilon -$ *best* responses. The second, which we refer to as an $\varepsilon -$ *predicting* QRE, insists on best response behavior but relaxes the requirement that players have the correct beliefs, and instead requires them to have approximately correct beliefs.

Given any two measures q and q$'$ on a metric space X, define

$$\| q - q' \| = Sup_{D \subseteq X} \mid q - q' \mid, \tag{5}$$

where the supremum is over Borel measureable subsets of X.

Definition 2.2. *Fix any $\varepsilon \geq 0$ and any normal form game with shocks $\Gamma = < I, A, E, r, \{u^i\}_{i \in I} >$. Also fix a distribution $\pi \in \mathcal{A}$.*

(a) π is an ε Quantal Response Equilibrium ($\varepsilon-$ QRE) for Γ if for some profile $\varepsilon - optimal$ action rules of players, α_ε, $\| \pi - \Delta(\alpha_\varepsilon, \pi) \| \leq \varepsilon$.

(b) π is an $\varepsilon - optimal$ Quantal Response Equilibrium ($\varepsilon - O$ QRE) for Γ if for some profile of $\varepsilon - optimal$ action rules of players, α_ε , $\pi = \Delta(\alpha_\varepsilon, \pi)$.

(c) π is an $\varepsilon - predicting$ Quantal Response Equilibrium ($\varepsilon - P$ QRE) for Γ if for some profile of best-response rules for players, α_*, $\| \pi - \Delta(\alpha_*, \pi) \| \leq \varepsilon$.

In the above definitions π represents the beliefs of players and $\Delta(\alpha, \pi)$ is the true ex ante play, generated by the action rules $\alpha = \{\alpha^i\}_{i \in I}$. In the definition of an $\varepsilon - O$ *QRE,* α is a collection of $\varepsilon -$ *optimal* action rules, and beliefs equal actual ex ante play. In the definition of an $\varepsilon - P$ QRE, α is a collection of best response rules, and, further, beliefs and actual play differ by at most ε. It should be clear that $\varepsilon - optimal$

$QRE's$ and $\varepsilon - predicting$ QRE's are $\varepsilon - QRE's$. Further, when $\varepsilon = 0$, $\varepsilon - optimal$ $QRE's$, $\varepsilon - predicting$ QRE's and $\varepsilon - QRE's$ become QRE's.

The above discussion talks about two ways in which $\varepsilon - QRE's$ are close to QRE's: via $\varepsilon - optimal$ rules and via $\varepsilon - close$ play. The proposition below provides yet another sense in which $\varepsilon - QRE$ are close to QRE's: to every $\varepsilon - QRE$ there exists a QRE which is close to it (as measured by the $\| \, . \, \|$ norm defined earlier). The proof appears in the appendix.

Proposition 2.1. *Fix a normal form game with shocks* $\Gamma =< I, A, E, r, \{u^i\}_{i \in I} >$. *Suppose the conditions of the Additive Shocks Model hold. Then* $\forall \delta > 0 \; \exists \bar{\varepsilon} > 0$ *such that* $\forall \varepsilon \leq \bar{\varepsilon}$, *for every* $\varepsilon - QRE \; \pi$, *there is a QRE* π^* *such that* $\| \pi - \pi^* \| \leq \delta$.

3 The Bayesian Repeated Game, B

3.1 The Model with "Types"

We will refer to the model of the earlier section as the model with "shocks" - the shocks are to the payoff functions. We now discuss the Bayesian repeated game model with "types." This is the framework that is used in most of the "rational" or Bayesian learning models mentioned in the introduction, and it provides a framework to impose the absolute continuity assumption we will use. In the next section we will discuss the dynamic extension of the shocks model of the previous section and indicate that it is a special case of this Bayesian model.

Just as before, there are I players, i=1,...,I playing a normal form game, this time repeatedly at dates n=1,2,... . Player i's action set is the finite set A^i, exactly the same at each date n, with $\mathcal{A}^i$ and $\mathcal{A}$ defined just as before. The set $H_N = A \times A \times A \times ... \times A$ (N-times) is the set of histories of length N. The set H_0 is the singleton set consisting of the null history, which we denote by h_0; and $H = \cup_{N=0}^{\infty} H_N$ is the set of all finite histories. Perfect recall assumed: when player i is choosing her date N+1 action she will know the date N history $h_N \in H_N$. The set of behavior strategies for player i is the set $F^i \equiv \{f^i : H \to \mathcal{A}^i\}$. We use the notation $F \equiv \prod_{i \in I} F^i$ and $F^{-i} \equiv \prod_{j \neq i} F^j$. The space F is endowed with the topology of weak convergence. Given any $f \in F$, $\phi(f)$ denotes the probability distribution over A^{∞} induced by f.

At the beginning of the game, period n=0, each player observes her type which is an element τ^i of her type space T^i (a complete and separable metric space). She does not observe the types of other players at any time. Define $T \equiv \prod_{i \in I} T^i$ and $T^{-i} = \prod_{j \neq i} T^j$. Let $\nu^i \in \wp(T^i)$ be the ex ante distribution governing the generation of player i's type $\tau^i \varepsilon T^i$. We suppose that the types of players are drawn independently, so that $\nu \equiv \otimes_{i=1}^{I} \nu^i$ is the distribution of the vectors of types in T. Let $u_n^i : A \times T^i \to \Re^1$ denote player i's date n utility function - in particular, the player i of type τ^i will receive at date n the utility of $u_n^i(a, \tau^i)$ when the vector of date n actions is $a \in A$. As before, we suppose that u_n^i is Borel measurable and bounded, with upper bound $\bar{u}$, and we normalize u_n^i so that it is non-negative. Define $U^i \equiv \{u_n^i\}_{n=1}^{\infty}$ and $u_n \equiv \{u_n^i\}_{i \in I}$.

We suppose that each player i forms two objects. (I) First, each player forms beliefs over the play of her opponents. We suppose that each player i believes that conditional on any date n-1 history, each of her opponents chooses a date n action independently of each other. We also suppose that player i's beliefs about her opponents is independent of her own realized type. Of course, player i may believe that player j's future actions will be influenced by i's own play. An implication of Kuhn's (1951) Theorem[1] is that such belief, b^i, can be represented as an element[2] of F^{-i}. We refer to such a b^i as a belief rule for player i. (II) Second, each player chooses a behavior rule, $f^i : T^i \rightarrow F^i$, indicating what behavior strategy she chooses as a function of her type[3].

If X is a cartesian product X=YxZ and $\eta \in \wp(X)$, we denote by $\text{Marg}_Y\ \eta$ the marginal of η on Y. Given any belief rule b^i and behavior rule f^i there is unique probability measure $\mu^i \in \wp(A^\infty \times T^i)$ such that

$$\text{MARG}_{T^i}\mu^i = \nu^i \text{ and} \tag{6}$$

$$\text{MARG}_{A^\infty}\mu^i(. \mid \tau^i) = \phi(f^i(\tau^i), b^i), \qquad \nu^i\text{-a.e. } \tau^i. \tag{7}$$

Any $\mu^i \in \wp(A^\infty \times T^i)$ which can be generated in the above manner from a tuple (ν^i, b^i, f^i) will be referred to as *an ex ante subjective belief* for player i. One may think of μ^i as the ex ante belief over $A^\infty \times T^i$ of a player i who believes her own types are generated by ν^i, who chooses a behavior rule f^i and whose belief about her opponents is represent by the belief rule b^i. Fix any collection of ex ante subjective beliefs for the players, $\{\mu^i\}_{i \in I}$, with associated vector $\{\nu^i, b^i, f^i\}_{i \in I}$ which generate it. Such a collection induces a unique "true" ex ante play, $\mu^* \in \wp(A^\infty \times T)$, defined to be the unique probability satisfying the following properties:

$$\text{MARG}_T\mu^* = \nu \text{ and} \tag{8}$$

$$\text{MARG}_{A^\infty}\mu^*(. \mid \tau) = \phi(\left\{f^i(\tau^i)\right\}_{i \in I}) \text{ for } \nu\text{-a.e. } \tau = \{\tau^i\}_{i \in I}. \tag{9}$$

Definition 3.1. A Bayesian game is any tuple $B = < I, A, T, v, \{U^i\}_{i \in I}, \{\mu^i\}_{i \in I}, \mu^* >$ with $\{\mu^i\}_{i \in I}$ a collection of ex ante subjective beliefs for players and μ^* the induced true distribution.

[1] See Kalai, E. and E. Lehrer (1993) for more on this.

[2] This construction is also equivalent, via a Kolmogorov Consistency theorem argument, to the formation of what Fudenberg and Kreps call an assessment rule: a sequence of *period-by-period* forecasts $\underline{\mu}^i = (\underline{\mu}_1^i, \underline{\mu}_2^i, ...)$ where $\underline{\mu}_n^i : H_n \times T^i \rightarrow \wp(A^{-i})$.

[3] Again, this is equivalent to the formation of what Fudenberg and Kreps (1993) call a behavior rule: a sequence of period-by-period plans $\underline{\phi}^i = (\underline{\phi}_1^i, \underline{\phi}_2^i, ...)$ where $\underline{\phi}_n^i : T^i \times H_n \rightarrow \mathcal{A}^{-i}$.

It should be stressed that the concept of a Bayesian game places no restrictions whatsoever on the beliefs or the actions of players. It should really be considered the language or framework for talking about beliefs and actions of players and how they evolve over time. In particular, the concept of a Bayesian game does not require players to be best-responding, to have common or correct beliefs or for any absolute continuity condition to hold.

The rational learning models of Jordan (1995), Kalai and Lehrer (1993), Nyarko (1998) and others can be embedded in the above framework. In those models the types are observed at date n=0 before the play of the game, and the payoff functions do not vary over time although they depend upon the realized types of players, so that

$$u_n^i(a, \tau^i) = u^i(a, \tau^i) \text{ for all n and for all } (a, \tau^i) \in A \times T^i. \tag{10}$$

3.2 Embedding "Shocks" Model into "Types" Model

In the literature using quantal response equilibria, use is made of the following dynamic extension of the static model with shocks presented in the earlier section. It is assumed that player i has a shock *process*, $\{e_n^i\}_{n=1}^{\infty} \in \mathcal{E}^{\infty}$. Before the date N decision is made player i would know the values of $\{e_n^i\}_{n=1}^{N}$. At date N the player will play the static game described in the earlier section but with the shock e_N^i replacing the e^i used before. We let $\Gamma^{\infty} = $ <I,A,$\{u^i\}_{i\in I}$,$\mathcal{E}^{\infty}$, r^{∞}> denote a generic repeated game with shocks. Recall that we used the notation Γ_{add} to represent the one period Additive Shocks Model. The following independence assumption will be used with the dynamic extension to Γ_{add}, which will be denoted by Γ_{add}^{∞}: *Player i's shock process,* $\{e_n^i\}_{n=1}^{\infty}$, *is independent and identically distributed (i.i.d) with common marginal distribution* r^i; *and* $\{e_n^i\}_{n=1}^{\infty}$ *is independent of* $\{e_n^j\}_{n=1}^{\infty}$ *for* $j \neq i$.

We now associate with any generic repeated game with shocks, Γ^{∞}, a Bayesian game B(Γ^{∞})=<I,A,T,v,$\{U^i\}_{i\in I}$, $\{\mu^i\}_{i\in I}$, μ^*> whose stochastic properties are exactly the same and which represents it in every meaningful manner. To this effect define player i's type space to be the sequence of realizations of the shocks $\tau^i = (e_1^i, e_2^i, ...) \in \prod_{n=1}^{\infty} \mathcal{E}^i$, and in particular set

$$T^i = \prod_{n=1}^{\infty} \mathcal{E}^i \quad \text{and} \quad \nu^i = \otimes_{n=1}^{\infty} r^i. \tag{11}$$

Next, define the utility function $U^i \equiv \{u_n^i\}_{n=1}^{\infty}$ by

$$u_n^i(a, \tau^i) = u^i(a, e_n^i) \qquad for \qquad \tau^i = (e_1^i, e_2^i, ...). \tag{12}$$

The only part of the repeated shocks model which does not seem to sit well with the Bayesian model with types is timing of when shocks and types are observed. In the shocks model the shocks are observed period-by-period. In the types model, the type is observed once and for all at date n=0 before the start of the game. We will however always assume that players are maximizing their stage game payoffs (see condition (BR) below), as opposed to some possibly discounted sum of infinite-horizon payoffs. Further, the shocks are assumed to be independent over time. In this case, it does not matter whether player i observes her shock period-by-period, or observes them all in one go at date 0.

The Bayesian game B of course also specifies ex ante subjective beliefs of players. The additive shocks game Γ^∞ is silent on beliefs, so we are free to include in B any subjective beliefs to effect the embedding of Γ^∞ into B. Once the additive shocks game is expanded to include a rule for belief formation, an ex ante subjective belief will emerge. If B and Γ^∞ share the same I and A, and if (11) and (12) hold, we shall say that B embeds Γ^∞ or that B is a Bayesian representation of Γ^∞.

4 Learning and Convergence to QRE: The Main Results

Throughout this section we shall consider as fixed some given Bayesian game $B = < I, A, T, v, \{U^i\}_{i\in I}, \{\mu^i\}_{i\in I}, \mu^* >$. Given a (Borel) probability P on some metric space X let Supp P denote the support of P. Player i is said to be (myopically) *best-responding* to her beliefs if for each of her possible types τ^i ($\nu^i - a.e.$) her behavior rule maximizes her expected utility at each date given her beliefs. In particular,

(BR) *(Best Responding)* For each i ∈ I, for $\nu^i - a.e.\ \tau^i$, and after each date N-1 history $h \in H_{N-1}$,

$$\text{Supp } f^i(\tau^i)(h) \subseteq Argmax_{a^i \in A^i} \int_{A^{-i}} u^i_N(a^i, a^{-i}, \tau_i) d\mu \left(. \mid \tau^i, h\right), \quad (13)$$

where the expectation above is with respect to the beliefs of player-type τ_i conditional on the history h and is taken over the set of possible date N actions, a^{-i}, of her opponents.

Given any two probability measures q and q′ on some metric space X, the measure q is said to be absolutely continuous with respect to q′ if for all Borel measurable subsets D⊆X, q(D)>0 implies that q′(D)>0. We then write q≪q′. We shall suppose that our fixed Bayesian game $B =< I, A, T, v, \{U^i\}_{i\in I}, \{\mu^i\}_{i\in I}, \mu^* >$ obeys the following the ex ante absolute continuity [4] condition.

[4] The absolute continuity assumption of Kalai and Lehrer (1993) is an ex post absolute continuity assumption. Ex Post absolute continuity implies but is not implied by ex ante absolute continuity. (See Nyarko (1998) for details.)

(AC) *(Absolute Continuity)* $\mu^* \ll \mu^i, \forall i \in I.$

Define

$$\pi_n^i \equiv M\arg_{A^i}\mu^*(\bullet \mid h_{n-1}), \qquad \pi_n \equiv \prod_{i\in I}\pi_n^i \text{ and } \pi_n^{-i} \equiv \prod_{j\neq i}\pi_n^j \tag{14}$$

$$b_n^i \equiv M\arg_{A^{-i}}\mu^i(\bullet \mid h_{n-1}); \tag{15}$$

$$\beta_n^i \equiv \pi_n^i \otimes b_n^i; \quad \text{and} \tag{16}$$

$$\nu_n \equiv M\arg_T \ \mu^*(\bullet \mid h_{n-1}) \tag{17}$$

where the first two marginals, π_n^i and b_n^i are taken over the set of date n actions conditional on the date n-1 history h_{n-1}. Specifically, π_n is the true ex ante date n play of the players conditional on history h_{n-1} but ex ante to the revelation of types; and b_n^i is player i's belief about her opponents conditional on date n history h_{n-1}, which by assumption is independent of i's type.

We will be providing results which say that beliefs become $\varepsilon-$ QRE. For this we need a distribution in $\mathcal{A}$ whereas $b_n^i \in \mathcal{A}^{-i}$. So we append to b_n^i the measure π_n^i, the true ex ante distribution of i's play, to obtain the tuple $\beta_n^i \equiv \pi_n^i \otimes b_n^i$ defined above, which represents beliefs of player i about all players including herself. Finally, the measure ν_n is the true marginal distribution over the type space T, again following the date n history h_{n-1}.

We now state our main results:

Theorem 4.1A. *Suppose that (BR) and (AC) hold.*

(i) (True play becomes ε-QRE) On all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall\varepsilon > 0, \exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$,

$$\pi_n \text{ is an } \ \varepsilon - QRE \ \ for \quad \Gamma_n \equiv < I, A, T, \nu_n, u_n > . \tag{18}$$

(ii) (Beliefs become ε-QRE) Suppose now that either that there are two players or, if there are more than two players, that any two players have the same beliefs about a third player. Along any sample path define $\hat{\pi}_n^i$ to be the common belief of players $j \neq i$ about player i's date n actions given the date n history, and define $\hat{\pi}_n \equiv \otimes_{i\in I}\hat{\pi}_n^i$. Then on all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall\varepsilon > 0, \exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$,

$$\hat{\pi}_n \text{ is an } \varepsilon - QRE \text{ for } \Gamma_n \equiv < I, A, T, \nu_n, u_n > . \tag{19}$$

The independence assumptions in the Additive Shocks Model imply that the $\Gamma_n \equiv < I, A, T, \nu_n, u_n >$ in the above theorem is independent of n and is what we referred to as Γ_{add}. Further, the Additive Shocks Model satisfies some continuity properties which enables us to strengthen part (ii) of the above by relaxing the common beliefs condition. In particular, we have the following:

Theorem 4.1B. *(The Additive Shocks Model) Let* Γ_{add} *be a static Additive Shocks Model and suppose that* Γ^{∞}_{add} is embedded in B. *Suppose that (BR) and (AC) hold.*

(i) (True play becomes ε-QRE) On all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall\varepsilon > 0$, $\exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$,

$$\pi_n \text{ is an } \quad \varepsilon - QRE \quad for \qquad \Gamma_{add}. \tag{20}$$

(ii) (Each player's beliefs become ε-QRE). On all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall\varepsilon > 0$, $\exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$ and for each $i \in I$,

$$\beta^i_n \text{ is an } \varepsilon - QRE \text{ for } \Gamma_{add}. \tag{21}$$

5 The Details

This section will present the details and slightly stronger versions of our main results above, as well as the intuition, proofs and some illustrative examples for those results. We will use the definitions and assumptions of section 4, and again we consider as fixed some given Bayesian game $B = < I, A, T, v, \{U^i\}_{i\in I}, \{\mu^i\}_{i\in I}, \mu^* >$.

5.1 Preliminaries

Our results will use the following well-known Blackwell and Dubins (1962) merging of opinions theorem:

Proposition 5.1. *(Blackwell and Dubins) Suppose the Bayesian game B obeys (AC). Then $\forall\ \varepsilon > 0$ and on each sample path excluding a set with μ^* zero probability, $\exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$,*

$$\parallel b^i_n - \pi^{-i}_n \parallel \leq \varepsilon. \tag{22}$$

5.2 Convergence of True Ex Ante Play

The following strengthens part (i) of our main result Theorem 4.1A, concluding that we have an $\varepsilon - O$ QRE rather than merely an $\varepsilon - QRE$.

Proposition 5.2A. *(True play becomes $\varepsilon - O$ QRE). Suppose that (BR) and (AC) hold. Then on all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall \varepsilon > 0$, $\exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N$,*

$$\pi_n \text{ is an } \quad \varepsilon - O\ QRE \ \ for \qquad \Gamma_n \equiv < I, A, T, \nu_n, u_n > . \tag{23}$$

The intuition behind this result is straightforward. Player i's actions are a best response to her beliefs b_n^i. From (22) it is easy to see that such actions are, for large n, an $\varepsilon -$ *optimal* response to beliefs π_n^{-i}. By construction such actions generate π_n^i, so π_n is an $\varepsilon - O$ QRE. The formal details of this proof, and all others, appear in the appendix.

The above result uses our $\varepsilon - O$ QRE concept. One may ask for the above result to be stated in terms the $\varepsilon - P$ QRE concept. The $\varepsilon - P$ QRE notion requires players to be best responding to their beliefs and for beliefs to be within ε of the truth. The problem however is that the best response map is not continuous in beliefs. Small changes in beliefs around a point of indifference may cause big changes in the best response. Hence, as the example below demonstrates, the above proposition is false when stated in terms of the $\varepsilon - P$ QRE concept. Immediately after the example we will indicate that for the Additive Shocks Model, this discontinuity disappears, so we are able to obtain an $\varepsilon - P$ QRE result for true play.

Example 5.1. *(True play not $\varepsilon - P$ QRE). Suppose there are two players A and B repeatedly playing the following normal form stage game:*

		B	
		Left	Right
A	Top	1,0	0,0
	Bottom	0,0	1,0

Suppose that A has a type space $T_A = [-1, 1]^\infty$, and let ν^A be the infinite product of uniform distributions on $[-1, 1]$. Player B's type space is degenerate (or there is a single type of B). Given any type $\tau^A = (\tau_1^A, \tau_2^A, ...)$ Player A will choose at date n action TOP (resp. BOTTOM) if τ_n^A is positive (resp. non-positive). Player A's belief about B is that B will randomize with equal probability in an i.i.d manner between actions LEFT and RIGHT at each date. The truth, however, is that Player B chooses actions LEFT and RIGHT at date n with probabilities 0.5+δ_n and 0.5-δ_n, *where $\delta_n > 0$ for all n and $\delta_n \to 0$ as $n \to \infty$ (in particular, Player B chooses LEFT with a vanishingly tiny bit more probability than RIGHT). Player B knows the ex ante probability, (1/2,1/2), with which A chooses her actions. Notice that both players are best responding to their beliefs. Also note that if δ_n converges to zero sufficiently fast then the ex ante absolute continuity assumption can be shown to hold. The true ex ante play, however, is the vector $(\pi_n^A, \pi_n^B) = ((1/2, 1/2), (0.5+\delta_n, 0.5-\delta_n))$. The best response for Player A against $\pi_n^B = (0.5+\delta_n, 0.5-\delta_n)$ is the pure action TOP. In particular, if α^A is a best-response rule for player A, then $\Delta^A(\alpha^A(\pi_n^B, .)) = (1, 0)$. Hence $\| \pi_n^A - \Delta^A(\alpha^A(\pi_n^B, .)) \| = \| (1/2, 1/2) - (0, 1) \| = 1/2$. The distribution (π_n^A, π_n^B) is therefore not an $\varepsilon - P$ QRE for any $\varepsilon < 1/2$.*

We now show that the discontinuity implicit in the above example does not obtain in the Additive Shocks Model so we may conclude that true play is eventually an $\varepsilon - P$ QRE. This result is therefore a stronger version of part (i) of our second main result, Theorem 4.1B for the Additive Shocks Model, obtaining $\varepsilon - P$ QRE as rather than merely $\varepsilon-$ QRE.

Proposition 5.2B. *(The Additive Shocks Model) Let Γ_{add} be a static Additive Shocks Model and suppose that Γ_{add}^{∞} is embedded in B. Suppose that (BR) and (AC) hold. Then on all sample paths excluding possibly a set with zero μ^*-probability the following is true: $\forall \varepsilon > 0$, $\exists N$ (which may depend upon ε and the sample path) such that $\forall n \geq N, \pi_n$ is an $\varepsilon - P$ QRE for Γ_{add}.*

5.3 Convergence of Beliefs

The proposition below indicates we may get a convergence result to $\varepsilon - P$ QRE's for the general (non-additive shocks) model if we state the result in terms of convergence of beliefs as opposed to convergence of true ex ante play. This next proposition assumes players start from the same beliefs about others. In particular, any two players have the same beliefs about the play of a third player. The proposition below is therefore a stronger version of part (ii) of our main result Theorem 4.1A, obtaining an $\varepsilon - P$ QRE rather than merely an $\varepsilon-$ QRE.

Proposition 5.3A *(Beliefs becomes $\varepsilon - P$ QRE). Suppose that (BR) and (AC) hold. Suppose either that there are two players or, if there are more than two players, that any two players have the same beliefs about a third player. Along any sample path define $\hat{\pi}_n^i$ to be the common belief of players $j \neq i$ about player i's date n actions* *given the date n history, and define* $\hat{\pi}_n \equiv \otimes_{i \in I} \hat{\pi}_n^i$. *Then on all sample paths excluding possibly a set with zero* μ^*-*probability the following is true:* $\forall \varepsilon > 0$, $\exists N$ *(which may depend upon* ε *and the sample path) such that* $\forall n \geq N$, $\hat{\pi}_n$ *is an* $\varepsilon - P$ *QRE for* $\Gamma_n \equiv <I, A, T, \nu_n, u_n>$.

Because of the continuity in the Additive Shocks Model even more is true: each agent's individual beliefs are $\varepsilon - P$ QRE's. Hence, with the Additive Shocks Model we do not have to impose the common beliefs assumption used in the previous proposition. Notice that this result is a strengthening of part (ii) of our second main result, Theorem 4.1B for the Additive Shocks Model, where we obtain an $\varepsilon - P$ QRE rather than merely an $\varepsilon-$ QRE.

Proposition 5.3B. *(The Additive Shocks Model) Let* Γ_{add} *be a static Additive Shocks Model and suppose that* Γ_{add}^{∞} *is embedded in B. Suppose that (BR) and (AC) hold. Then on all sample paths excluding possibly a set with zero* μ^*-*probability the following is true:* $\forall \varepsilon > 0$, $\exists N$ *(which may depend upon* ε *and the sample path) such that* $\forall n \geq N$ *and for each* $i \in I$, β_n^i *is an* $\varepsilon - P$ *QRE for* Γ_{add}.

The results of this sub-section on beliefs will in general not be correct if we replace $\varepsilon - P$ QRE with $\varepsilon - O$ QRE. In particular, even under the common beliefs assumption used in the earlier proposition, beliefs do not become $\varepsilon - O$ QRE. The latter requires for all $\varepsilon > 0$ the existence of a behavior rule profile $\{\alpha^i\}_{i \in I}$ which is $\varepsilon-$*optimal* against the beliefs $\hat{\pi}_n$ and is such that an exact relation, $\hat{\pi}_n^i = \Delta^i(\alpha^i(\hat{\pi}_n^{-i}, .))$, holds for all dates n sufficiently large. In particular, for all large n, the beliefs that others have about player i, $\hat{\pi}_n^i$, should be exactly equal to true play of player i, $\Delta^i(\alpha^i(\hat{\pi}_n^{-i}, .))$. Although learning occurs over time, in most cases there will never be a date n for which beliefs equal actual play exactly. A simple example is provided below to illustrate this[5].

[5] For an easy example not using the model with i.i.d shocks consider the following. Suppose Player A has three actions. Suppose she plays her first action all the time. Player B assigns equal probability to two kinds of behavior by A. The first kind of behavior of A is what A is truly doing - playing her first action at all dates. The second kind of action is an equal randomization at each date of each one of her other actions in an i.i.d manner. Hence at each date Player B's beliefs will assign positive probability to an action of A other than her first. Set the stage game payoffs so that if A plays any of her other actions she gets a large disutility relative to choosing her first action. Then the only $\varepsilon -$ *optimal* behavior rule for A is to choose the first action at each date, which will result in play which is not the same as B's beliefs at any date.

Example 5.2. *(Beliefs not $\varepsilon - O$ QRE's.) Suppose there are two players, A and B, and suppose B has only one action and one type. Suppose A has two actions TOP and BOTTOM. The payoff to BOTTOM is zero. The date n payoff to TOP equals A's date n type, τ_n^A, which is uniformly distributed on the union of the two intervals $\mathcal{E}^A$ =[-5,-10] $\cup$ [5,10]. Fix any $\varepsilon < 5$. It should be clear that the only $\varepsilon -$ optimal action rule for player A is the same as the best-response rule, α_*^A: choose TOP if $\tau_n^A \in$ [5,10] and to choose BOTTOM otherwise. In particular, any $\varepsilon -$ optimal action rule will result in A choosing each of her two actions with equal ex ante probability at each date. Suppose that B knows that A's shocks at each date are generated in an i.i.d manner from $\mathcal{E}^A$ (as they really are). But B does not know whether the common marginal distribution is uniform on $\mathcal{E}^A$ or is some other distribution, $\hat{\rho}^A$ say. Player B assigns equal probability to each of these two. Player B will, over time, learn the true distribution of play of A, π_n^A (which is equal probability on each action at each date in an i.i.d manner). This learning will however will in general be gradual. On no finite date n will B's belief about A's true ex ante play, b_n^A, be exactly equal to A's true ex ante play, π_n^A. In particular, B's beliefs will never be an $\varepsilon - O$ QRE.*

6 Conclusion

We have provided conditions under which repeated play and learning causes movement toward a quantal response equilibrium. This provides some foundation for the use of such equilibria.

The results of this paper have implicitly used many independence assumptions. In particular, it has been assumed that (a) each player believes that others are choosing actions independently of each other and their types are drawn independently of each other and (b) each player chooses actions independently of others and (c) in the Additive Shocks Model it is assumed that the shocks occur in an i.i.d manner. Parts (a) and (b) can be relaxed along the lines of Nyarko (1994) and similar convergence results may be obtained if we generalize the concept of a quantal response equilibrium to some kind of correlated quantal response equilibrium.

We have argued elsewhere[6] that the appropriate concept of convergence in many repeated games with imperfect information is in terms of the ex ante (i.e., not conditioning on realized types) or in terms of beliefs. The true play of players in the repeated shocks model are typically pure actions since these are conditional on realized values of the shock process. The quantal response equilibrium, on the other hand, is usually an equilibrium in beliefs - it represents what players believe about the ex ante distribution of play, ex ante to receipt of shocks to the utility. The ex post play is not an equilibrium; the ex ante is. The convergence results obtained in

[6] See Nyarko (1998) .

this paper are of precisely that kind[7], and show the convergence of ex ante play to quantal response equilibria.

7 Appendix: The Proofs

7.1 Proof of Proposition 2.1.

Proof. Define, for each $\delta > 0$, $\mathcal{O}(\delta)$ to be the set of all distributions which are within distance δ of some QRE:

$$\mathcal{O}(\delta) = \{\pi \in \mathcal{A} : \parallel \pi - \pi^* \parallel \leq \delta \text{ for some QRE } \pi^*\}. \tag{24}$$

Suppose that the proposition is false. Then $\exists \bar{\delta} > 0$ such that $\forall k = 1, 2, 3, \ldots$, there exists a $\frac{1}{k}- QRE$, π_k, such that $\pi_k \notin \mathcal{O}\left(\bar{\delta}\right)$. Since $\mathcal{A}$ is compact, we know that $\{\pi_k\}_{k=1}^{\infty}$ has a convergent subsequence, so without loss of generality we may suppose that π_k converges as k$\rightarrow \infty$ to some $\pi_\infty \in \mathcal{A}$. For each k, since π_k is a $\frac{1}{k}- QRE$, there exists a profile of $\frac{1}{k} - optimal$ action rules, α_k, such that

$$\parallel \pi_k - \Delta(\alpha_k, \pi_k) \parallel \leq \frac{1}{k}. \tag{25}$$

Fix any m=1,2,... , and recall that α_* denotes an profile of best-response rules. From the claim below we may conclude that

$$\parallel \Delta(\alpha_k, \pi_k) - \Delta(\alpha_*, \pi_k) \parallel \leq \frac{1}{m} \text{ for all k sufficiently large.} \tag{26}$$

From the continuity of $\Delta(\alpha_*, \pi)$ in π we may conclude that

$$\parallel \Delta(\alpha_*, \pi_k) - \Delta(\alpha_*, \pi_\infty) \parallel \leq \frac{1}{m} \text{ for all k sufficiently large.} \tag{27}$$

The above three inequalities imply that

$$\parallel \pi_k - \Delta(\alpha_*, \pi_\infty) \parallel \leq \frac{1}{k} + \frac{2}{m} \text{ for all k sufficiently large.} \tag{28}$$

Taking k and then m to infinity implies that $\pi_\infty = \Delta(\alpha_*, \pi_\infty)$, so π_∞ is a QRE. Since $\pi_k \rightarrow \pi_\infty$, this is a contradiction to the fact that $\pi_k \notin \mathcal{O}\left(\bar{\delta}\right)$ for all k, and proves the proposition.

CLAIM: Fix any $\bar{\pi}$ and any $\delta > 0$. *Then* $\exists \bar{\varepsilon} > 0$ and ρ>0 such that such that for all $\varepsilon \leq \bar{\varepsilon}$, if α_ε is a profile of $\varepsilon - optimal$ action rules and $\pi \in \mathcal{A}$ is such that $\parallel \pi - \bar{\pi} \parallel \leq \rho$, then

$$\parallel \Delta(\alpha_\varepsilon, \pi) - \Delta(\alpha_*, \pi) \parallel \leq \delta. \tag{29}$$

[7] Fudenberg and Kreps (1993) obtain convergence in the same ex ante sense.

PROOF OF CLAIM: Fix any i $\in I$. Order the action set A^i and write $A^i = \{1, 2, ..., \#A^i\}$. Let $\mathcal{U}^i \equiv \Re^{\#A^i}$. Given any belief of i about her opponents, the set $\mathcal{U}^i$ is the set of possible vectors of expected utilities from i's actions - in particular, any vector x=$\{x_k\}_{k=1}^{\#A^i} \in \mathcal{U}^i$ is a vector of expected utilities with x$_k$ the expected utility to action k. Given any belief $\pi^{-i} \in \mathcal{A}^{-i}$ that player i could have about her opponents, we define P^i (π^{-i}) to be the distribution over the expected utilities of player i, ex ante to observation of i's shock e^i: for any set C $\subseteq \mathcal{U}^i$,

$$P^i\left(\pi^{-i}\right)(C) \equiv r\left(\left\{e^i \in \mathcal{E}^i : \left\{\int_{A^{-i}} u^i\left(k, a^{-i}, e^i\right) d\pi^{-i}\right\}_{k=1}^{\#A^i} \in C\right\}\right). \tag{30}$$

We now define for each action K and each $\varepsilon > 0$, the set $\mathcal{D}_K^\varepsilon$ to be the set of expected utility vectors of player i such that action K is could be chosen under an $\varepsilon - optimal$ rule while some other action could be chosen by a best-response rule: in particular, define

$$\mathcal{D}_K^\varepsilon \equiv \left\{x = \{x_k\}_{k=1}^{\#A^i} \in \mathcal{U}^i : Max\{x_k\}_{k\neq K} \geq x_K \geq Max\{x_k\}_{k\neq K} - \varepsilon\right\} \text{ and} \tag{31}$$

$$\mathcal{D}^\varepsilon \equiv \cup_{K=1}^{\#A^i} \mathcal{D}_K^\varepsilon \tag{32}$$

The set $\mathcal{D}^\varepsilon$ above is the only set of expected utility vectors where the best-reponse rule and $\varepsilon - optimal$ rule could prescribe different actions. Hence,

$$\| \Delta^i(\alpha_\varepsilon^i, \pi^{-i}) - \Delta^i(\alpha_*^i, \pi^{-i}) \| \leq P^i\left(\pi^{-i}\right)(\mathcal{D}^\varepsilon). \tag{33}$$

Fix any $\bar{\pi}$ and any $\delta > 0$. The set $\mathcal{D}^0$ has zero Lebesgue measure, so under the Additive Shocks Model assumptions $P^i(\bar{\pi}^{-i})(\mathcal{D}^0) = 0$. Since $\mathcal{D}^\varepsilon \downarrow \mathcal{D}^0$, we conclude that for some $\bar{\varepsilon} > 0$,

$$P^i\left(\bar{\pi}^{-i}\right)(\mathcal{D}^{\bar{\varepsilon}}) \leq \frac{\delta}{2}. \tag{34}$$

Under the Additive Shocks Model assumptions one may show that for fixed C $\subseteq \mathcal{U}^i$, $P^i(\pi^{-i})(C)$ is continuous in π^{-i} and in particular $\exists \rho > 0$ such that

$$\| P^i\left(\pi^{-i}\right)(\mathcal{D}^{\bar{\varepsilon}}) - P^i\left(\bar{\pi}^{-i}\right)(\mathcal{D}^{\bar{\varepsilon}}) \| \leq \frac{\delta}{2} \text{ for all } \pi^{-i} \in \mathcal{A}^{-i} \text{ such that } \| \pi^{-i} - \bar{\pi}^{-i} \| \leq \rho. \tag{35}$$

Combining (33), (34) and (35)and noting that $\mathcal{D}^\varepsilon$ is decreasing in ε implies that

$$\| \Delta^i(\alpha_\varepsilon^i, \pi^{-i}) - \Delta^i(\alpha_*^i, \pi^{-i}) \| \leq \delta \text{ for all } \pi \text{ such that } \| \pi^{-i} - \bar{\pi}^{-i} \| \leq \rho \text{ and } \forall \varepsilon \leq \bar{\varepsilon}. \tag{36}$$

From this we obtain the claim. ■

7.2 Proof of Proposition 5.2A.

Fix any $\varepsilon > 0$ and any sample path, excluding those for which (22) does not hold. Let $\bar{u}$ be the bound on the utility function u_i^n. Let N be large enough so that (22) implies that for all n$\geq N$,

$$\| b_n^i - \pi_n^{-i} \| \leq \frac{\varepsilon}{\bar{u}\,(\#A)}, \tag{37}$$

where #A is the cardinality of the action set A. Fix any such n.

Recall that $\{\alpha_*^i\}_{i\in I}$ are the best-response action rules. Define the action rule α^i by

$$\alpha^i\left(b^i,\tau^i\right) = \begin{cases} \alpha_*^i(b_n^i,\tau^i) \text{ if } b^i = \pi_n^{-i} \\ \text{any arbitrary value if } b^i \neq \pi_n^{-i} \end{cases} \tag{38}$$

so that

$$\alpha^i\left(\pi_n^{-i},\tau^i\right) = \alpha_*^i(b_n^i,\tau^i) \text{ for } \nu - a.e. \ \ \tau^i \tag{39}$$

Since players are best responding to their beliefs, $\pi_n^i = \Delta^i(\alpha_*^i(b_n^i,.))$ so (39) implies that

$$\pi_n^i = \Delta^i(\alpha^i(\pi_n^{-i},.)). \tag{40}$$

Next, best-response behavior also implies that $\forall a^i \in A^i$

$$\int_{A_{-i}} u_n^i(a^i,a^{-i},\tau^i)db_n^i \leq \int_{A_{-i}} u_n^i(\alpha_*^i(b_n^i,\tau^i),a^{-i},\tau^i)db_n^i, \tag{41}$$

which from (37),

$$\leq \int_{A_{-i}} [u_n^i(\alpha_*^i(b_n^i,\tau^i),a^{-i},\tau^i)d\pi_n^{-i} + \varepsilon \tag{42}$$

which in turn, from (39) ,

$$\leq \int_{A_{-i}} [u_n^i(\alpha^i\left(\pi_n^{-i},\tau^i\right),a^{-i},\tau^i)d\pi_n^{-i} + \varepsilon. \tag{43}$$

We have therefore shown that α^i is an $\varepsilon - \textit{optimal}$ rule at π_n^{-i}. Combining this with (40) proves the proposition.

7.3 Proof of Proposition 5.2B

Let $\{\alpha_*^i\}_{i\in I}$ be a collection of best-response rules. Players are best responding to their beliefs so π_n^i=$\Delta^i(\alpha_*^i(\mathrm{b}_n^i,.))$. From (22)we know that $\|\mathrm{b}_n^i$-π_n^{-i} $\|\rightarrow 0$. In the the Additive Shocks Model, however, $\Delta^i(\alpha_*^i(b^i,.))$ is continuous in b^i (see Fudenberg and Kreps (1993, Lemma 7.3)). Hence, $\|$ π_n^i-$\Delta^i(\alpha_*^i(\pi_n^{-i},.))\|\rightarrow 0$. From this the proposition follows immediately.

7.4 Proof of Proposition 5.3A

Fix any sample path, any $\varepsilon > 0$, and any date n. By assumption agents are best responding to their beliefs, so $\pi_n^i = \Delta^i(\alpha_*^i(\hat{\pi}_n^{-i}, .))$. Eq. (22) therefore implies that $\| \hat{\pi}_n^i - \Delta^i(\alpha_*^i(\hat{\pi}_n^{-i}, .)) \| \leq \varepsilon$ for all i in I. This proves the proposition.

7.5 Proof of Proposition 5.3B

Fix any i$\in I$. Define $\zeta_n = \beta_n^i$ and $\zeta_n^j = Marg_{A^j}\zeta_n$ for each j$\in I$. We seek to show that ζ_n is an $\varepsilon - P$ QRE $for\ \Gamma_{add}$. Fix any j$\in I$. From (22) as n becomes large

$$\| \zeta_n^j - \pi_n^j \| \rightarrow 0. \tag{44}$$

By assumption players are best responding, so for all n,

$$\pi_n^j = \Delta^j(\alpha_*^j(b_n^j, .)). \tag{45}$$

From (22) as n becomes large $\| \beta_n^j - \beta_n^i \| \rightarrow 0$ so $\| b_n^j - \zeta_n^{-j} \| \rightarrow 0$. In the Additive Shocks Model, $\Delta^i(\alpha_*^i(b^i, .))$ is continuous in b^i (see Fudenberg and Kreps (1993, Lemma 7.3)), so as n becomes large

$$\| \Delta^j(\alpha_*^j(b_n^j, .)) - \Delta^j(\alpha_*^j(\zeta_n^{-i}, .)) \| \rightarrow 0. \tag{46}$$

From $(44) - (46)$ we conclude that as n becomes large,

$$\| \zeta_n^j - \Delta^j(\alpha_*^j(\zeta_n^{-i}, .)) \| \rightarrow 0. \tag{47}$$

This implies that ζ_n is an $\varepsilon - P$ QRE $for\ \Gamma_{add}$.

References

[1] Blackwell, D. and L. Dubins (1963):"Merging of Opinions with Increasing Information," Annals of Mathematical Statistics, 38, 882-886.

[2] Fudenberg, D. and D. Kreps, (1993) "Learning Mixed Equilibria", Games and Economic Behavior", 5,pp. 320–367.

[3] Jackson, M. and E. Kalai (1997): "Social Learning in Recurring Games, "Games and Economic Behavior", 21,pp. 102-134.

[4] Jordan, J. (1997):"Bayesian Learning in Games: A NonBayesian Perspective," in The Dynamics of Norms, C. Bicchieri, R. Jeffery and B. Skyrms (eds.), Cambridge University Press, Cambridge, 1997.

[5] Jordan, J. (1995),"Bayesian Learning in Repeated Games", Games and Economic Behavior, 9,pp. 8–20.

[6] Kalai E. and E.Lehrer (1993) "Rational Learning Leads to Nash Equilibrium", Econometrica, 61, pp. 1019-46.

[7] McKelvey, R. and T. Palfrey (1995),"Quantal Response Equilibrium for Normal Form Games", Games and Economic Behavior,10, pp. 6–38.

[8] Nyarko, Y. (1994): "Bayesian Learning in Leads to Correlated Equilibria in Normal Form Games",Economic Theory, 11, pp. 643–656.

[9] Nyarko, Y. (1998): "Bayesian Learning and Convergence to Nash equilibrium Without Common Priors", Economic Theory, 11, pp. 643–656.