Discussion Paper No. 1151

RECURRING BULLIES, TREMBLING AND LEARNING

by

Matthew Jackson and Ehud Kalai

Northwestern University

April 1995

Recurring Bullies, Trembling and Learning[1]

Matthew Jackson and Ehud Kalai

Kellogg Graduate School of Management, Northwestern University, Evanston, IL 60208-2009

April 1995

Abstract. In a recurring game, a stage game is played consecutively by different groups of players, with each group receiving information about the play of earlier groups. Starting with uncertainty about the distribution of types in the population, late groups may learn to play a correct Bayesian equilibrium, as if they know the type distribution.

This paper concentrates on Selten's Chain Store game and the Kreps, Milgrom, Roberts, Wilson phenomenon, where a small perceived inaccuracy about the type distribution can drastically alter the equilibrium behavior. It presents sufficient conditions that prevent this phenomenon from persisting in a recurring setting.

Keywords. Recurring Game, Social Learning, Chain Store Paradox

1 Introduction

In a recurring game, a stage game is sequentially played by different groups of players. Each group, before its turn, receives information about the social history consisting of past plays of earlier groups. Game theorists have studied such recurring situations using various dynamics like fictitious play, last period best response, and random matching, since Nash in his dissertation (1950).

Of interest here is a Bayesian version of a recurring game, where each stage game player is uncertain about the types of his or her opponents. Observed social histories are used by players to update beliefs about the unknown distribution of types in the population. With time, players' behavior converges to that of the Bayesian equilibrium of the stage game, as if the true distribution of types in the population were known.

In a recent paper (Jackson and Kalai, 1995a) we present a general model of recurring games and sufficient conditions that yield such social convergence. Our purpose in this chapter is to apply this approach to Selten's (1978) chain-store games and study the implications on the phenomenon illustrated by Kreps and Wilson (1982) and Milgrom and Roberts (1982) (KMRW for short, also sometimes called the "gang of four"). A specific question is whether this phenomenon

[1]The authors thank the California Institute of Technology and the Sherman Fairchild Foundation for their generous hospitality while doing this research. They also thank Eric van Damme for comments on an early draft and the National Science Foundation for financial support under grants SBR-9223156 and SBR-9223338.

(described in more detail below) can persist in a recurring setting. To better fit the recurring setting we replace Selten's chain-store story with a strategically equivalent story about a bully terrorizing a finite group of individuals, the challengers. The rest of the introduction presents a verbal description of our model and conclusions. A broader survey of the related literature can be found in Jackson and Kalai (1995a).

The "challenge the bully" stage game (the bully game, for short) is played by a bully and L challengers. It consists of L preordered episodes each played by the bully and one challenger. In each episode the designated challenger chooses whether to challenge the bully or not, and if challenged, the bully chooses whether to fight or not. The bully prefers not to be challenged, but if challenged he would rather not fight. The challenger prefers challenging a nonfighting bully, but if the bully fights he would rather not challenge. (The single episode payoffs are represented in Figure 1 of the next section.) Both bully and challenger know the outcome of all previous episodes before making their decisions.

Recall that Selten's paradox is that backward induction applied to this game dictates that all challengers challenge and the bully never fights. Yet, when L is large, common sense suggests that the bully may choose to fight in early episodes in order to build a reputation so that challengers will not challenge.[2] KMRW point out that the backward induction result depends on the complete information assumption. The resolution offered by KMRW is to replace the above game by a Harsanyi (1967–68) Bayesian game in which there is a small commonly known prior probability that the bully is "irrational," e.g., he prefers fighting to any other episode outcome. In such a modified game the more reasonable behavior is obtained as the unique sequential equilibrium outcome, even when the realized bully is rational. Thus the phenomenon illustrated by KMRW is that in a world with a rational bully, a small uncertainty about the rationality of the bully can drastically change the equilibrium outcomes, and in particular, induce the rational bully to fight.

It is not obvious whether this phenomenon will persist if this situation recurs. In other words, if new bullies are always born rational, but challengers are uncertain about this fact, would bullying behavior persist? On the one hand, it seems that statistical updating will lead observers (i.e., players in later rounds) through backward induction to the recognition that the bullies are rational and with it will come challenging and not fighting. On the other hand, a violation of a rational act by a bully that wishes to appear irrational, is more convincing when done in view of such learning. Thus, one is left with the question of whether such learning will lead the eventual discovery of the rationality of the population and thus the unraveling of the KMRW phenomenon, or whether the actions of the players will slow the learning so that the rationality of the

[2] There are, of course, Nash equilibria of this type; however, they do not survive backward induction.

population is never learned.

To meaningfully address this question, we extend the KMRW game to a "doubly Bayesian" recurring game with two "priors." One prior, τ, is a fixed probablity distribution according to which a bully-type is drawn before each stage game. We assume, however, that this type-generating distribution is not known to the players. They assume that it was selected before the start of the entire recurring game according to a known probability distribution Γ over a set of possible type generating distributions.[3]

This gives rise to what we call an *uncertain recurring game* whose extensive form game can be described as follows. In period 0 nature randomly selects a type generating probability distribution τ from a set of possible ones, according to the prior probabilities $\Gamma(\tau)$. Players know that a selection was made according to Γ, but no information is given about the realized τ. For period 1, one bully and L challengers are selected. The bully is randomly assigned a type according to the fixed (unknown) type-generating distribution τ. Only the bully is informed of its realized type and the $L+1$ players proceed to play the L episodes of the bully game. Their realized play path becomes publicly known. The game recurs, following the same procedure. In each period $t > 1$ a new set of challengers and a bully are selected. Again, the bully is privately assigned its type according to the same unknown τ and the players proceed to play the L episodes of the bully game, with the play path publicly revealed.

To study the recurring version of the KMRW phenomenon we will focus the above described game, with a realiztion of τ that selects rational bullies with probability one, indicating a world in which bullies are really rational.

Note that the above separation to two distributions is necessary since a restriction to a single prior τ, representing both actual type-generation and social beliefs, will force one of the following misrepresentations of the phenomenon. If the single τ assigns probability one to a rational bully, then there would be no social uncertainty about it. On the other hand, if the single τ assigns positive probability to irrational bully types, then in the recurring setting with probability one some irrational types will be realized, leading to the KMRW equilibrium, but in a world which truly has irrational bullies. In either situation, there would be no social learning. That is, observing past stage games would tell a player nothing new about the distribution of types in the population.

Our recuring game definition of a type is also significantly generalized. A

[3]This approach may be viewed as the multi-arm-bandit "payoff" learning model adapted to the Bayesian repeated game "type" learning literature. [See Aumann and Maschler (1967) and followers, and the more closely related Jordan (1991).] The utility maximizing bandit player starts with a prior distribution over his set of possible payoff generating distributions. Ours is a multi-person version, with uncertainty over the distribution generating types, modeled through a prior over a set of possible distributions. A major difference between our model of learning in recurring games and the literature on learning in repeated games is that in our setting players are attempting to learn about the distribution of types that they will face in their stage game, while in a repeated game players learn about the actual opponents they repeatedly face.

type's preferences can depend on the entire social history and not just on his own stage game. So, for example, we can model a "macho bully" whose utility of fighting increases after social histories with many earlier fighting bullies. Or in a multi-generational setting, a bully may prefer to behave like his ancestors.

We refer to the Nash equilibria of the above *uncertain recurring game* as *uncertain Bayesian equilibria*. In playing an uncertain Bayesian equilibrium, expected utility maximizers perform Bayesian updating of the prior Γ to obtain updated posteriors over the set of type generating distributions. As a result, the strategies of each stage game constitute a Bayesian equilibrium of the stage game relative to the perceived distribution over types induced by the updated prior, but not necessarily relative to the true (realized) type generating distribution.

This discrepancy partly disappears however, after long social histories. First, as a consequence of the martingale convergence theorem, the updated prior will converge (almost everywhere). This means that players' learning disappears in the limit. Second, the updated distribution will be empirically correct in the limit. This means that players' learning induced predictions concerning the play path will match the distribution over play paths induced by the true (realized) distribution over types. This second result can be obtained almost directly from a learning result in Kalai and Lehrer (1993). It means that in late games, any discrepancy between the true and the learning induced type generating distributions cannot be detected, even with sophisticated statistical tests.

The fact that players learn to predict the play path with arbitrary precision, does not necessarily imply that they have learned the true type generating distribution τ or the true distribution of strategies (including off the equilibrium path behavior).[4] We present an example of challengers who never challenge, because their initial beliefs assign high probability to fighting bullies, even though the realized bullies would never fight if challenged. No learning occurs because no challenging ever occurs. Moreover, in the example a single tremble by a single challenger will reveal the non fighting nature of the bullies and would cause the entire equilibrium to collapse.

In order to overcome this difficulty, we introduce trembles à la Selten (1975, 1983). The effect of introducing the trembles is to ensure that the social history leads to informative learning.[5] For an exogenously given small positive probability q, we assume that every player chooses strategies that assign probabilities of at least q to each one of his or her actions at every information set.

Assuming that players play an uncertain Bayesian equilibrium with trem-

[4]This phenomenon also arises in non-recurring situations, such as those where players play a long extensive form or a repeated game. For detailed analyses of this in such settings, see Battigalli, Gilli and Molinari (1992), Fudenberg and Kreps (1988), Fudenberg and Levine (1993), and Kalai and Lehrer (1993).

[5]See the concluding remarks of Fudenberg and Kreps (1995) for the suggestion of a similar approach to ensure learning and convergence to Nash equilibrium in a model with myopic long lived players.

bling, we obtain the following conclusions for the case where the realized distribution results in only rational bullies: With probability one, in late stage games, the challengers always challenge, and the bullies never fight. These rules are followed with the exception of occasional trembles. This result is obtained regardless of the initial beliefs of the society, provided those initial beliefs assign some positive probability to the distribution which selects only rational bullies.

Thus the introduction of the trembles in the recurring model leads one to eventually play the trembling hand perfect equilibrium (defined with respect to the agent normal form). It is important to remark, however, that the trembles are not working directly (as in the definition of trembling hand perfection), but rather indirectly through the learning that they ensure. This is evident since in early stages, the equilibrium outcomes with trembles can still be that rational bullies fight and challengers not challenge.

We wish to remark that the general message of our results is not simply the eventual decay of the KMRW equilibrium. For instance, if the realized type generating distribution is truly selecting some "irrational" fighting bullies, then this would also eventually be learned and the eventual convergence would be to the KMRW equilibrium - even if initial beliefs placed an arbitrarily high (but not exclusive) weight on rational bullies. Thus the more general understanding of our results is that in a recurring game, given some randomness to induce learning, equilibrium behavior will eventually converge to that as if players were playing an equilibrium knowing the underlying type generating distribution.[6] If players are truly rational, and the game is an extensive form, then this will lead to eventual approximate play of a perfect equilibrium.

2 The Recurring Bully Game

The *stage game* is a *bully game* consisting of one bully player, b, and L challengers, $(c_i)_{i=1,\dots,L}$. In sequential episodes $i = 1, \dots, L$, challenger c_i decides whether to challenge (C) the bully or refrain (R) from doing so. If the challenger does challenge, then the bully has to decide whether to acquiesce (A) or fight (F). The corresponding episode payoffs, to the challenger and bully, are given by the extensive form game pictured in Figure 1.

Each player is informed of the outcomes of earlier episodes before making a decision. The payoffs of challengers are determined according to their episode payoffs while the payoff to the bully is the sum of his or her L episode payoffs.

Formally, a *play path* of such a bully game is described by a vector $p = (X_1, \dots, X_L)$ with each X_i, being in the set $\{(R),\ (C, A),\ (C, F)\}$, describing the outcome of the i^{th} episode. A *partial play path* is such a vector $(X_1, \dots, X_l)$, but with $0 \leq l < L$. A (behavioral) strategy σ_i of challenger c_i consists of a probability distribution over the actions R and C for every $i - 1$ long partial play path $(X_1, \dots, X_{i-1})$. A bully strategy, η, chooses a probability distribution

[6] See Jackson and Kalai (1995ab) for general results along these lines.

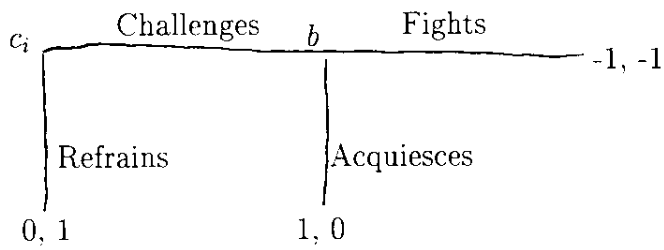

Figure 1

over the actions A and F for every partial play path of every length. A vector of strategies $((\sigma_i), \eta)$ induces a probability distribution over the set of play paths and defines expected payoffs for all players in the obvious unique way.

The *recurring bully game* is played in periods $t = 1, 2, \ldots$. In each period t a new group of players (bully and challengers) are selected to play the stage game. Moreover, before choosing their strategies, they are informed of the play paths of all earlier groups.

A *social history* h of length t is a vector of play paths $h = (p^1, \ldots, p^t)$. The empty history $\emptyset$ is the only history of length 0, and H denotes the set of all finite length histories. The players of the recurring bully game are denoted by b^h and c_i^h, where for every fixed h, b^h and c_i^h describe the bully and the challenger to play the i-th episode of the stage bully game after history of h. Thus, as the notation suggests, the players b^h and c_i^h know the social history h that led up to their game. Strategies in the recurring bully game are denoted by η^h and σ_i^h, with the obvious interpretations.

3 Games with Unknown Bully Types

Alternative bully types describe different utility functions that a bully may have in playing a bully game. There is a countable set Θ of possible bully types. For each type $\theta \in \Theta$ there is a function $u_\theta(h, p)$ describing the payoff to bully b^h of the play path p, when he or she is of type θ.

An example of a type is the one already defined whose payoff from any path (regardless of the social history) is the sum of the episode payoffs described earlier (in Figure 1). We refer to these payoffs as the "rational" ones and to this type as the *rational type*, denoted by r.

An example of an "irrational" bully, motivated by KMRW, is one with payoff $k > 0$ for every "strong path" p in which he or she never acquiesces, and $-k$ for weak paths p, exhibiting some acquiescing actions. We keep the terms "rational" and "irrational" in quotes, since "irrational" bullies are supposed to be maximizing their expected payoffs. It is simply that the "irrational" payoffs do not match those of the game in Figure 1. Of course, this is simply a modeling convenience as any behvior can always be made to be expected utility maximizing, if one can arbitrarily manipulate the payoff function.

More complex "irrational" types can condition their payoffs on the observed social history. For example in the type just described we can replace k by k^h, taking on large values after social histories h consisting only of tough play paths and small values for social histories h with weak bullies. This may represent an ego maniac bully who compares himself to earlier ones.

Probability distributions over the countable set of possible bully types, Θ, are referred to as *type-generating distributions.* Every such distribution τ defines a *recurring Bayesian bully game* as follows. After every social history h, nature chooses a bully type θ^h, according to the distribution τ (independently of all earlier choices in the game). The bully b^h is informed of his or her realized type θ^h, before choosing his strategy. All other players are only aware of the process and know the distribution τ by which θ^h was selected.

Strategies in the Bayesian recurring bully game are denoted by (η_θ^h) and (σ_i^h), with η_θ^h denoting the strategy of the bully b^h when he or she is of type θ, and σ_i^h is a challenger strategy as before.

Notice that a Bayesian recurring bully game, with a type generating distribution τ, induces after every history h a *Bayesian bully game* of reputation, similar to KMRW, with bully types payoffs given by $u_\theta(p, h)$ and a prior τ on possible types θ.

We wish to model, however, Bayesian recurring bully games with uncertainty about the type generating distribution. It is such uncertainty that introduces learning from the social history. An *uncertain Bayesian recurring bully game* (uncertain recurring game for short) is played as follows. In a 0-time move, nature selects a type generating distribution τ according to a prior probability distribution Γ. Without being informed of the realized τ, the players proceed to play the τ Bayesian recurring bully game. We assume that the exogenously given distribution Γ has a countable support ($\Gamma(\tau) > 0$ for at most countable many τ's) and that it and the structure of the game are commonly known to all players.[7] Players update Γ based on observed histories and the strategies that they believe to be governing play. This updated distribution induces a distribution over types, which is the basis for the Bayesian stage game that they play.

Notice that an uncertain recurring bully game may be thought of as a "doubly Bayesian" game since we use a prior to select a distribution that serves as the priors of the stage games to come. The strategies of the uncertain recurring game are the same as the ones of the Bayesian game since the information transmitted and feasible actions are identical in both games. However, as will

[7] see Jackson and Kalai (1995a) for discussion of extensions to allow for players' prior beliefs to be type dependent.

be seen in the sequel, expected utility computations and Bayesian updating are more substantial in the uncertain recurring game.

An *uncertain Bayesian equilibrium* is a vector of strategies $((\eta_\theta^h)), (\sigma_i^h))$ which are best reply to each other in the extensive form description of the uncertain recurring game defined by the prior Γ.

4 Social Learning

In this section, we consider a given uncertain Bayesian recurring bully game with a prior Γ and fixed strategies $((\eta_\theta^h), (\sigma_i^h))$. We first clarify some probabilistic issues.

A fully described *outcome* of the uncertain recurring game is a sequence of the type $\tau, \theta^1, p^1, \theta^2, p^2, \ldots$. Such an outcome generates progressive socially-observed histories $h^1, h^2 \ldots$ with $h^t = (p^1, \ldots, p^t)$. To describe the probability distribution on the set of all outcomes it suffices to define consistent probabilities for all initial segments of outcomes. We do so inductively by defining $P(\tau) = \Gamma(\tau)$ and

$$P(\tau, \theta^1, p^1, \ldots, \theta^{t+1}, p^{t+1}) = P(\tau, \theta^1, p^1, \ldots, \theta^t, p^t)\tau(\theta^{t+1})\mu(p^{t+1})$$

where $\mu(p^{t+1})$ is the probability of p^{t+1} under the strategies $((\eta_\theta^h), (\sigma_i^h))$ with $h = h^t$ and $\theta = \theta^{t+1}$.

Players, observing only the buildup of histories $h^1, h^2, \ldots$, do not know the chosen type generating distribution τ, but they can generate posterior probability distribution over it using the initial distribution and conditioning on current histories. Thus, their posteriors $\Gamma^0, \Gamma^1, \ldots$, are defined for every type generating distribution $\bar{\tau}$ by $\Gamma^0(\bar{\tau}) = \Gamma(\bar{\tau})$ and $\Gamma^t(\bar{\tau}) = P(\bar{\tau} \mid h^t)$.

Each such distribution over type generating distributions induces a direct distribution over types. Thus we have *updated posterior distributions on the next period type* denoted by $\gamma^0, \gamma^1, \ldots$, with $\gamma^t(\theta) = \sum_{\bar{\tau}} \bar{\tau}(\theta)\Gamma^t(\bar{\tau})$. Although we supress the notation, both Γ^t and γ^t are history dependent.

After every history h^t the bully, b^{h^t}, and the challengers, $(c_i^{h^t})$, play a Bayesian bully game. The bully knows his or her realized type, θ^{t+1}, and it is commonly assumed that he or she was drawn according to the distribution γ^t. If the original strategies $((\eta_\theta^h), (\sigma_i^h))$ constitute an uncertain Bayesian equilibrium of the recurring game, then $((\eta_\theta^h), (\sigma_i^h))$ with $h = h^t$ and $\theta = \theta^{t+1}$ constitute a Bayesian equilibrium of the stage Bayesian game with the prior γ^t.

Of course, the assumed prior, γ^t, against which individual challenger optimal strategies are chosen, is "wrong", since the real prior, by which θ^{t+1} is chosen, is the unknown τ. In Harsanyi's (1967–68) definition of a Bayesian equilibrium this presents no problem. The strategies, with the commonly known assumption that the prior is γ^t, are formally a Bayesian equilibrium. But in a recurring setting, when types are repeatedly drawn, a discrepancy between a real prior and an assumed prior may lead to statistical contradictions. These statistical discrepencies will disappear as players learn from observed histories.

To describe the effects of learning we proceed in two steps. Our first proposition states that players' updated distributions (over distributions) converge almost surely. Effectively, after some random time, players stop learning.

Proposition 1 *For almost every outcome, $\gamma^t(\theta)$ converges to a limit $\gamma(\theta)$ uniformly for all types $\theta \in \Theta$.*

Proposition 1 follows from the martingale convergence theorem. We refer readers to Jackson and Kalai (1995a) for a proof.

Our next proposition states that when players stop learning they have in fact learned all that they could, and so they are arbitrarily correctly predicting the play path after some random time. Proposition 2 can be proven using Proposition 1, or can be seen more directly as a consequence of Theorem 3 in Kalai and Lehrer (1993).[8]

With the fixed stage game strategies, let μ be the probability distribution induced on the stage game play paths by the real prior, τ, and let $\tilde{\mu}$ be the one induced by the assumed prior, γ^t. If for every play path $\mu(p^{t+1}) = \tilde{\mu}(p^{t+1})$ then no observable contradictions, even statistical ones can arise. We refer to the strategies and the assumed prior γ^t as an *empirically correct Bayesian equilibrium* whenever $\mu = \tilde{\mu}$. If for some $\epsilon > 0$ we have $| \mu(p^{t+1}) - \tilde{\mu}(p^{t+1}) | \leq \epsilon$ for all play paths p^{t+1}, we refer to it as an *empirically ϵ-correct Bayesian equilibrium.*

Proposition 2 *In an uncertain Bayesian equilibrium of an uncertain recurring bully game, for almost every outcome and for every $\epsilon > 0$, there exists a time T, such that γ^t, together with the induced $t+1$ period strategies, constitute an empirically ϵ-correct Bayesian equilibrium of the stage game for all $t \geq T$.*

The fact that late period stage-game players play an empirically ϵ-correct Bayesian equilibrium relative to their updated type generating distribution γ^t does not mean that they play a Bayesian equilibrium relative to the correct distribution τ. It only means that players' predictions concerning the play path are approximately correct. Players may be mistaken concerning off the equilibrium path behavior. This is illustrated in the following example.

Example

Consider three types of a bully, $r, f,$ and a, described by their episode payoffs as follows. The *rational* type, r, has the original extensive game payoffs described earlier. The *fighting* type, f, has a payoff 1 for fighting and 0 for any

[8]Theorem 3 in Kalai and Lehrer (1993) states that Bayesian updating of beliefs containing a "grain of truth" must eventually lead to correct predictions. The restriction in the current model, to a countable set of type generating distributions, implies that for almost every outcome $\Gamma(\tau) > 0$, which implies that the beliefs contain a grain of truth.

other outcome (replacing the payoffs in Figure 1), while the *acquiescing* type, a, has a payoff of 1 for acquiescing and 0 for any other outcome. Let ρ, ψ and α, be the type generating distributions that select respectively with probability one the types r, f and a.

Suppose the prior Γ assigns high probability to a world with fighting bullies, i.e., $\Gamma(\psi)$ is high, and only a small probability to α and to ρ. Suppose the real type generating distribution was selected to be α, i.e. only acquiescing bullies are generated. The strategies with all challengers refrain, a and r type bullies always acquiesce, and f type bullies always fight, constitute an uncertain Bayesian equilibrium. This is so because the initial stage game prior γ^0 assigns high probability to a fighting bully and to refrain from challenging is therefore rational. But since there is no challenge in the first period the updated posterior on bully types remains unchanged, i.e. $\gamma^1 = \gamma^0$, the same logic applies to the second period, and so on. Moreover, the induced stage game Bayesian equilibria are empirically accurate. It is clear, however, that these strategies do not induce a Bayesian equilibrium of the stage game relative to the real prior α. In other words, if the challengers knew that their bully is drawn with probability one to be of the a type, then they would challenge.

5 Trembling Rational Players

The uncertain Bayesian equilibrium in the above example is highly unstable. If at any time, in the infinite play of the recurring game, a challenger trembles and challenges, the whole equilibrium will collapse after observing that the bully does not fight. The equilibrium is able to survive only because society never learns what bullies would do if challenged.

If a small amount of noise, in the form of trembles by players, is introduced, then more learning will occur and equilibria such as the one in the example above will eventually be overturned. In short, a small amount of imperfection will lead players to learn behavior at all nodes in the tree. While there are several ways to model such trembles, we follow Selten's (1975) approach by restricting the set of strategies that a player can choose.

Let q be a positive small number describing the probability of minimal trembles. The *uncertain Bayesian recurring bully game* with *q-trembling* is obtained from the usual uncertain recurring bully game by restricting the players to the choice of behavior strategies that assign probability greater or equal to q to every action available in each information set.

Our next observation is that under q-trembling, if a Bayesian equilibrium is empirically ϵ-correct, then it must also be approximately correct for all conditional probabilities in the game. For example, the probability that the bully fights the i^{th} challenger, conditional on the event that the i^{th} challenger challenges, must be similar when computed by the learning induced distribution γ and by the true (realized) distribution τ.

To see this point, let $\tilde{\mu}$ be the distribution induced on play paths through the learning induced distribution γ, and let μ be the one obtained from the true (realized) distribution τ. The fact that $\mid \mu(p) - \tilde{\mu}(p) \mid \leq \epsilon$ for any play path p implies that this can be made (by starting with a smaller ϵ) true for any event (i.e. a set of play paths) in the game. The q-trembling property implies that there is a small positive number s such that the probability of any non empty event in the game is at least s. The conditional probability of event E_1 given E_2 is $P(E_1$ and $E_2)/P(E_2)$. Since we have all denominators (over all possible E_2's) being uniformly bounded from zero, by making ϵ sufficiently small we can make all such ratios when computed by μ or by $\tilde{\mu}$ be within any given δ of each other. Thus we obtain the following.

Define $\tilde{\mu}$ to be *strongly empirically ϵ-correct* if for any two events E_1, and E_2 the conditional probabilities of E_1 given E_2, computed by μ and by $\tilde{\mu}$ are within ϵ of each other.

Lemma 3 *Consider any Bayesian equilibrium $((\eta_\theta), (\sigma_i))$ of a q-trembling Bayesian bully game with an assumed type generating distribution γ and a true type generating distribution τ. For every $\epsilon > 0$ there is a $\delta > 0$ such that if the equilibrium is empirically δ-correct it must be strongly empirically ϵ-correct.*

Notice that by combining the result of Proposition 2 with the above lemma, we can conclude that for almost every outcome of an uncertain Bayesian equilibrium of the recurring game, after a sufficiently long time, stage game players must be playing a strongly empirically ϵ-correct Bayesian equilibrium.

We consider now the special case, of an uncertain Bayesian equilibrium of the recurring game, where the true (realized) type generating distribution, ρ, generates the rational type r with probability one. The prior Γ over type generating distributions can be arbitrary, as long as $\Gamma(\rho) > 0$.

For a game with q-trembling, a bully stage strategy η is *essentially acquiescing* if its probability of fighting at every information set is the minimally possible level q, i.e. is only due to trembling. Similarly, a challenger i strategy is *essentially challenging* if the probability of refraining is q. We can now state the following result.

Proposition 4 *Consider an uncertain Bayesian equilibrium of an uncertain recurring bully game with q-trembles ($q > 0$). If the realized type generating distribution places weight one on rational players, then for small enough q ($1/2 > q$) and for almost every outcome there is a finite time T such that the t-period stage game strategies must be essentially acquiescing and essentially challenging for all $t \geq T$.*

Given the previous propositions and lemma, the conclusion of this proposition is straightforward. For every $\epsilon > 0$ there is a sufficiently large T so that all the later stage games are played by strongly empirically ϵ-correct Bayesian equilibrium. So all we have to observe is that for sufficiently small ϵ, a strongly empirically ϵ-correct Bayesian equilibrium must consist of essentially acquiescing and essentially challenging strategies. This is done by a (backward) induction argument, outlined as follows.

In the last episode, following any play path, a rational bully must acquiesce with the highest possible conditional probability given the (positive probability) event that the last challenger challenges. Therefore the rational bully must acquiesce with probability $1 - q$ in the last episode. Thus, after time T (as defined in Lemma 3) the last episode challenger's assessed probability of the bully acquiescing in that episode is at least $1 - q - \epsilon$. Therefore, for sufficiently small ϵ and $q < 1/2$ the last challenger's unique best response is to challenge, and so he or she challenges with probability $1 - q$. Since this is true for any play path leading to the final episode, the assessed probability of challenge in the final episode is at least $1 - q - \epsilon$, indepedent of the play path leading to that episode. It follows that a rational bully in episode $L - 1$, will have a unique best response of acquiescing in that episode. The $L - 1$ period challenger, assessing this to be the case with probability at least $1 - q - \epsilon$, will essentially challenge, and so on.

6 Concluding Remarks

The KMRW phenomenon extends to a general folk theorem for finitely repeated games with incomplete information, as shown by Fudenberg and Maskin (1986). The failing of the phenomenon in a recurring setting has parallel implications for this folk theorem. Jackson and Kalai (1995ab) contains general results that have direct implications on this question. Again, we should emphasize that our results imply that under certain conditions players will learn to play as if they knew the realized type generating distribution. To the extent that there exists a true diversity of types in the population, this is learned and players will play accordingly. Thus, the variety of equilibrium outcomes allowed for by the folk theorem can still be realized, but will require a true diversity of types, rather than just a perceived one, in order to survive in a recurring setting.

Stronger, and even more striking, versions of the KMRW phenomenon might be possible in the manner described by Aumann (1992). This would involve higher order misconceptions on the part of players. For example, replace the KMRW situation, where challengers are uncertain about whether the bully is rational or fight loving, by a situation where all challengers know that the bully is rational, but are uncertain about whether other challengers also know it. If, for instance, challengers believe that other challengers believe that there may be a fighting bully, then the KMRW results might be extended. This situation is incorporated into our model by using the type space to allow an explicit description of the beliefs a player holds about the beliefs of other players (and choosing τ's and Γ to reflect this uncertainty). Proposition 4 covers these cases and thus, if bullies are born rational, then this extended version of the KMRW equilibrium would also unravel in a recurring setting.

Getting back to Selten's paradox, it seems to become more severe in the recurring setting. A bully in later stages of the game who fights the first challenger can be explained as having trembled, and thus is not perceived as being irrational. Moreover, in the equilibrium play in late enough stages, even a bully who fights the first few challengers will be explained as having trembled several times, as this likelihood is larger than the alternative explanation of the bully being irrational. Although in late enough play it is relatively more likely that this behavior is due to trembles rather than irrationality, both of these events were very unlikely to start with. Thus, it may be that the challenger would prefer to doubt the model altogether (or believe that the bully has done so), rather than to ascribe probabilities according to it. Thus we are back at Selten's paradox.

Let us close with two comments relating to "technical" assumptions that we have maintained in our analysis. The restriction in this chapter to countably many types is convenient for mathematical exposition. The generalization to an uncountably infinite set (and an uncoutably infinite set of possible type generating distributions) requires additional conditions, in particular to assure that Proposition 2 extends. Lehrer and Smorodinsky (1994) offer general conditions which are useful in this direction.

One strong assumption we have made is that players start from a common prior over priors. This is not necessary for the results. The content of Propositions 1 and 2 can be applied to situations where each type of player has their own beliefs. The equilibrium convergence result then needs to be modified, since players' beliefs may converge at different rates. When these convergence rates are not uniform, then the conclusions are stated relative to a set of types which receives a probability arbitrarily close to one (under the realized distribution).

7 References

Aumann, R. J. (1992), "Irrationality in Game Theory," in *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, edited by P. Dasgupta, D. Gale, O. Hart, and E. Maskin, MIT Press, Cambridge MA, 214–227.

Aumann, R. J. and M. Maschler (1967), "Repeated Games with Incomplete Information: A Survey of Recent Results," *Mathematica*, ST - 116, Ch. III, 287–403.

Battigalli, P., M. Gilli, and M.C. Molinari (1992) "Learning and Convergence to Equilibrium in Repeated Strategic Interactions: An Introductory Survey," *Richerche Economiche*, **96**, 335–378.

Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting and Incomplete Information," *Econometrica*, **54**, 533-554.

Fudenberg, D. and D. Kreps (1988), "A Theory of Learning, Experimentation and Equilibrium in Games," mimeo: Stanford University.

Fudenberg, D. and D. Kreps (1995), "Learning in Extensive Form Games: Self Confirming Equilibrium," *Games and Economic Behavior*, **8**, 20–55.

Fudenberg, D. and D. Levine (1993), "Steady State Learning and Nash Equilibrium," *Econometrica*, **61**, 547–573.

Harsanyi, J. (1967–68), "Games of Incomplete Information Played by Bayesian Players (Parts I, II, and III)", *Management Science*, **14**, 159-182; 320-334; 486–503.

Jackson, M. and E. Kalai (1995a), "Social Learning in Recurring Games", mimeo, California Institute of Technology.

Jackson, M. and E. Kalai (1995b), "Learning to Play Perfectly in Recurring Extensive Form Games", mimeo, California Institute of Technology.

Jordan, J. (1991), "Bayesian Learning in Normal Form Games," *Games and Economic Behavior*, **3**, 60–81.

Kalai, E. and E. Lehrer (1993), "Rational Learning Leads to Nash Equilibrium", *Econometrica*, **61**, 1019-1045.

Kreps, D. and R. Wilson (1982), "Reputation and Imperfect Information", *Journal of Economic Theory*, **27**, 253-279.

Lehrer, E. and R. Smorodinsky (1994), "Repeated Large Games with Incomplete Information", Tel-Aviv University.

Milgrom, P. and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory*, **27**, 280-312.

Nash, J. (1950), "Non-Cooperative Games", PhD Thesis, Mathematics Department, Princeton University.

Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory*, **4**, 25-55.

Selten, R. (1978), "The Chain-Store Paradox", *Theory and Decision*, **9**, 127-159.

Selten, R. (1983), "Evolutionary Stability in Extensive 2-Person Games," *Mathematical Social Sciences*, **5**, 269–363.