REPUTATION WITH LONG RUN PLAYERS AND IMPERFECT OBSERVATION

ALP E. ATAKAN AND MEHMET EKMEKCI

Abstract.

Previous work shows that reputation results may fail in repeated games between two long-run players with equal discount factors. We restrict attention to an infinitely repeated game where two players with equal discount factors play a stage game where actions are imperfectly observed and the first mover advantage is maximal. The set of types for player 1 is taken as any countable subset of the set of all finite automaton. In this context a one sided reputation result is provided. If player 1 is a Stackelberg type with positive probability and player 2's actions are imperfectly observed, then player 1 receives the highest individually rational payoff in all Bayes Nash equilibria, as agents become patient.

Keywords: Repeated Games, Reputation, Equal Discount Factor, Long-run Players, Imperfect Observability, Complicated Types, Finite Automaton
JEL Classification Numbers: C73, D83.

1. Introduction and Related Literature

This paper proves a reputation result for a class of two-player repeated games where the first mover advantage is maximal (maximal Stackelberg games). The players have equal discount factors. Our main result shows that if there is incomplete information about the type of player 1 and player 2's payoffs are imperfectly observed, then player 1 receives the highest possible payoff, in any Bayes Nash equilibrium, as the discount factors converge to one.

Date: This Draft, July, 2008.

We would like to thank Christoph Kuzmics for useful discussions about complicated commitment types.

Reputation effects were first established for finitely repeated games by Kreps and Wilson (1982) and Milgrom and Roberts (1982); and extended to infinitely repeated games by Fudenberg and Levine (1989). Fudenberg and Levine (1989, 1992) focused on repeated games where one long-run player faces a sequence of short-run players. The authors considered an incomplete information version of the repeated game where the long-run player is a "Stackelberg" (or commitment) type who always plays the "Stackelberg" (or commitment) action. They proved a *one-sided reputation result* that shows if the long-run player is sufficiently patient, then the long-run player's equilibrium payoff is at least as large as what the player could get by publicly committing to the Stackelberg action. However, most reputation results in the literature are for repeated games where one of the players is a short-run player (as in Fudenberg and Levine (1989, 1992)); or for repeated games where the player building the reputation is infinitely more patient than his rival and so the rival is essentially a short-run player, at the limit (for example, see Schmidt (1993b) or Celantani, Fudenberg, Levine, and Pesendorfer (1996)).

Previous research has also shown that reputation results may fail when long-run players with equal discount factors play an infinitely repeated game in which the stage game is a simultaneous-move game in normal form. In particular, one-sided reputation results obtain only if the stage game is a game of *strictly conflicting interest* (Cripps, Dekel, and Pesendorfer (2005)), or the commitment action is a dominant action in the stage (Chan (2000)).[1] For all other simultaneous-move games, folk theorems by Chan (2000) and Cripps and Thomas (1997) have shown that any individually rational and feasible payoff can be sustained in perfect equilibria of the infinitely repeated game, if the players are sufficiently patient.

A *Stackelberg action* is a player's optimal stage game strategy, if the player could publicly commit to this strategy; and a *Stackelberg type* is a commitment type that only plays the Stackelberg action.[2] In a *maximal-Stackelberg game* the Stackelberg action delivers player

[1]A game has strictly conflicting interests (Chan (2000)) if a best reply to the commitment action of player 1 yields the best feasible and individually rational payoff for player 1 and the minimax for player 2.

[2]In our set-up a pure stage game action is a pure strategy profile of the extensive-form stage game.

1 his highest payoff, that is part of an individually rational payoff profile, whenever player 2 best responds.

This paper focuses on simultaneous-move maximal Stackelberg stage games. Player 2's actions are imperfectly observed by player 1 while player 1's actions are perfectly observed. The set of commitment types for player 1 is taken as any countable subset of the set of finite automaton that includes a Stackelberg type. In this framework we prove a one-sided reputation result. Specifically, we show that player 1 receives a payoff arbitrarily close to his highest individually rational payoff for sufficiently high discount factor, for any initial probability distribution over the set of commitment types. This result is in sharp contrast with previous literature – with arbitrarily small imperfect observability player 1 can guarantee a high payoff whereas with perfect observability Cripps and Thomas (1997) and Chan (2000) demonstrate a folk theorem, for the same class of games. Also, the result presented here improves upon the reputation result of Cripps, Dekel, and Pesendorfer (2005) for *strictly conflicting interest* games since our result allows for an arbitrary subset of commitment types

The paper proceeds as follows: section 2 describes the model and discusses some examples that satisfy our assumptions; section 3 presents the main result; and section 4 concludes.

2. The Model

Let Γ denote the stage game. The set of players in the game is $I = \{1, 2\}$. The set of pure stage game actions, a_i, for player i is denoted A_i and the set of mixed stage game strategies α_i is $\Delta(A_i)$. Player 2's actions are not observed by player 1. Player 1's actions are perfectly observed by player 2. Let the finite set Y_2 denote the set of publicly observed outcomes of player $2's$ actions. Each action profile $a_2 \in A_2$ induces a probability distribution over the publicly observed outcomes Y_2. The set of all public outcomes $Y = A_1 \times Y_2$ includes information on player 1's actions, since these are perfectly observed, and the outcomes of player 2's actions. For a given action a_2, let $\pi_y(a_2)$ denote the probability of outcome $y \in Y_2$.

Assumption 1 (Uniform Imperfect Observability)**.** *For any a_2 and $a_2' \in A_2$, $supp(\pi(a_2)) = supp(\pi(a_2'))$.*

Assumption 1 implies that player 1 is never exactly sure about player 2's actions. The assumption, however, does not put any limits on the degree of imperfect observability. In fact, player 1's information can be arbitrarily close to perfect information.

Stage game payoffs $r_i(a_1, y_2)$ depend only on publicly observed outcomes. So, actions matter only in their influence over the distribution of outcomes. Let

$$g_i(a_1, a_2) = \sum_{y_2 \in Y_2} r_i(a_1, y_2)\pi_y(a_2)$$

and so the stage game Γ with action sets A_i and payoff function $g_i : A_1 \times A_2 \to \mathbb{R}$ is a standard normal form game. The minimax for player i, $\hat{g}_i = \min_{\alpha_j} \max_{\alpha_i} g_i(\alpha_i, \alpha_j)$. The set of feasible payoffs $F = co\{(g_1(a_1, a_2), g_2(a_1, a_2)) : (a_1, a_2) \in A_1 \times A_2\}$; and the set of feasible and individually rational payoffs $G = F \cap \{(g_1, g_2) : g_1 \geq \hat{g}_1, g_2 \geq \hat{g}_2\}$. Largest feasible and individually rational payoff for player i, $\bar{g}_i = \max\{g_i : (g_i, g_j) \in G\} = 1$. Also, let $M > \max_{g \in F} |g_i|$.

Assumption 2 (Maximal-Stackelberg Game for Player i)**.** *There exists $a_i^s \in A_i$ such that if a_j is a best response for player j to action a_i^s, then $g_i(a_i^s, a_j) = \bar{g}_i$. Also, for any two payoff vectors $g \in G$ and $g' \in G$, if $g_i = g_i' = \bar{g}_i$, then $g_j = g_j'$.*

Assumption 2 requires that there exists an action for player 1, denoted a_1^s, which gives player 1 his highest payoff that is part of an individually rational payoff profile whenever player 2 best replies to a_1^s. Also, the assumption requires that the payoff profile that delivers player 1 payoff $\bar{g}_1$ is unique.[3] This uniqueness requirement is satisfied generically by extensive-form games.

Definition 1 (Strictly Conflicting Interest for Player i)**.** *The game satisfies Assumption 2 for player i. Also, for all $(\bar{g}_i, g_j) \in G$, $g_j = \hat{g}_j$.*

[3]Under this Assumption 2 there can be more than one Stackelberg action for player 1 and player 2 may have multiple best responses to any of player 1's Stackelberg actions.

If Γ satisfies Assumption 2 for player 1 but is not a game of strictly conflicting interest for player 1, then there is a finite constant $\rho \geq 0$ such that

$$\left| \frac{g_2 - g_2(a_1^s, a_2^b)}{\bar{g}_1 - g_1} \right| \leq \rho \tag{1}$$

for any $(g_1, g_2) \in F$, where a_2^b denotes a best reply to a_1^s.[4] Also, if Γ is a game of strictly conflicting interest for player 1, then $g_2 - g_2(a_1^s, a_2^b) \leq \rho(\bar{g}_1 - g_1)$ for any $(g_1, g_2) \in F$.

In the repeated game Γ^∞, the stage game Γ is played in each of periods $t = 0, 1, 2, ...$ Players perfectly recall their own past actions and all past outcomes. Let $H_t^2 \triangleq (A_1 \times A_2 \times Y_2)^t$ denote the set of all partial histories, h_t^2 , that can be observed by player 2 before the start of period t. Likewise let $H_t^1 \triangleq (A_1 \times Y_2)^t$ denote the set of all partial histories that can be observed by player 1 before the start of period t. A behavior strategy is a function $\sigma_i : \bigcup_{t=0}^{\infty} H_t^i \to \Delta(A_i)$. A behavior strategy chooses a mixed stage game strategy given the partial history h_t^i. Players discount payoffs using their discount factor δ. The players' continuation payoffs in the repeated game are given by the normalized discounted sum of the continuation stage-game payoffs

$$u_i(h_{-t}) = (1 - \delta) \sum_{k=t}^{\infty} \delta^{k-t} g((a_i)_k, (a_j)_k)$$

where h_{-t} denotes the continuation play in periods $t, ...\infty$ in some history $h = \{(a_1, a_2)_0, ..., (a_1, a_2)_k, ...\}$.

The set of possible types for player 1, denoted Ω, is a countable subset of the set of all *finite automatons*. A *finite automaton* (Q, q_0, f, τ) consists of a finite set of states Q, an initial state $q_0 \in Q$, an output function $f : Q \to \Delta(A_1)$ that assigns a (possibly mixed) action to each state, and a transition function $\tau : Y_2 \times A_1 \times Q \to Q$ that assigns a state to each outcome of the stage game. Specifically, the action chosen by an automaton in period t is determined by the output function f given the period t state (q_t) of the automaton. The evolution of states for the automaton is determined by the transition function τ.

---

[4]If Γ satisfies Assumption 2 but not Assumption 3 then $\bar{g}_1 = \max\{g_1 : (g_1, g_2) \in F\}$. Consequently, the Lipschitz condition given in Equation (1) holds for all $g \in F$ and not just for $g \in G$.

Let μ denote a probability measures over Ω with $\mu(\omega) > 0$ for each $\omega \in \Omega$. Before time 0 nature selects player 1 as a type ω with probability $\mu_i(\omega)$. Ω contains a normal type denoted ω_N. The normal type maximizes expected normalized discounted utility. Ω also contains a Stackelberg type denoted ω_s. A Stackelberg type for player i plays a_i^s in period 0 and also after any history where player i has played only a_1^s. Ω_s denotes the set of all Stackelberg types. Let $\Omega_- = \Omega \backslash (\{\omega_N\} \cup \Omega_s)$. In words, Ω_- is the set of types other than the Stackelberg types and the normal type.

A belief for player 2, $\mu : \bigcup_{t=0}^{\infty} H_t^2 \to [0,1]$ assigns a probability measure over Ω after each partial history h_t^2. Let $\sigma_1(\omega)$ denote the repeated game strategy used by type ω, i.e., $\sigma_1(\omega_N)$ is the strategy used by player 1's normal type and $\sigma_1(s)$ the strategy used by player 1's Stackelberg type. A strategy profile, $\sigma = (\{\sigma_1(\omega)\}_{\omega \in \Omega}, \sigma_2)$, is a list of the strategies used by each type of player 1, as well as, the strategy of player 2. Player 2's expected continuation utility, following a partial history h_t, given that player 2 uses the strategy profile σ

$$U_2(\sigma|h_t) = \mu(\omega_N|h_t)\mathbb{E}_\sigma[u_2(h_{-t})|h_t] + \mu(\Omega_s|h_t)\mathbb{E}_{(\sigma_1(s),\sigma_2)}[u_2(h_{-t})|h_t] + \sum_{\omega \in \Omega_-} \mu_j(\omega|h_t)\mathbb{E}_{(\sigma_1(\omega),\sigma_2)}[u_2(h_{-t})|h_t]$$

where $\mathbb{E}_{(\sigma_j,\sigma_i)}[u_i(h_{-t})|h_t]$ denotes the expectation over continuation histories h_{-t} generated by (σ_j, σ_i) given that h_t has occurred. Also, player 1's normal type's payoff following a partial history h_t, $U_1(\sigma|h_t) = \mathbb{E}_\sigma[u_1(h_{-t})|h_t]$.

The repeated game where the initial probability over Ω is μ and the discount factor is δ is denoted $\Gamma^\infty(\mu, \delta)$. The analysis in the paper focuses on Bayes-Nash of the game of incomplete information $\Gamma^\infty(\mu, \delta)$. In equilibrium, beliefs are obtained, where possible, using Bayes' rule given $\mu(\cdot|h_0) = \mu(\cdot)$ and conditioning on players' equilibrium strategies. If $\mu(\Omega_s) > 0$, then belief $\mu(\cdot|h_t)$ is well defined after any history where player 1 has played a_1^s in each period.

2.1. **Examples.** The following are examples of Generic Stackelberg Games. These examples are not games of strictly conflicting interest, and also, the players do not have strictly dominant actions in these games. Consequently, if in any of these stage games player 2's actions are perfectly observed, then Folk theorems apply implying that there is no scope for reputation effects. (See Cripps and Thomas (1997) for Example 1, and Chan (2000) for Examples 1 through 4). That is, in any of these games, any individually rational payoff profile can be sustained as a perfect equilibrium of the repeated game, for sufficiently high discount factor and sufficiently small probability of facing a commitment type. In marked contrast, if the player 2's actions are imperfectly observed (i.e., if the games satisfy Assumption 1), then Theorem 1 provides a one-sided reputation result.

2.1.1. *Common Interest Games.* Consider the following simplified common interest game as depicted in Figure 1.

		Player 2 L	Player 2 R
P1	U	$1,1$	$0,0$
	D	$\epsilon,0$	$0,0$

FIGURE 1. Common Interest Game ($\epsilon < 1$).

Assume that there is a (possibly small) probability that player 1 is a commitment type that always plays the Stackelberg action (U) and player 2's actions are imperfectly observed. In particular, suppose that after player 2 plays L (or R) with probability ϵ the observation is R (or L) and with probability $1-\epsilon$ the correct action is observed. Then Theorem 1 implies that player 1 can guarantee a payoff arbitrarily close to 1 in any perfect equilibrium of the repeated game, for sufficiently high discount factors.

2.1.2. *Battle of the Sexes.* In the version of the battle of the sexes depicted in Figure 2 both players prefer action profiles on the diagonal of the payoff matrix but Player 1 prefers (U,L) while player 2 prefers (D,R). The minimax payoff for each player is $1/2$ and is attained by the action profile (M,M).

		Player 2 L	R	M
	U	$2,1$	$0,0$	$0,0$
P1	D	$0,0$	$1,2$	$0,0$
	M	$0,0$	$0,0$	$1/2,1/2$

Figure 2. Battle of the Sexes.

3. Main Reputation Result

The main result of the paper, Theorem 1, shows that Player 1 can guarantee a payoff arbitrarily close to his highest individually rational payoff for high enough discount factor. The set of possible types for player 1 is taken as any countable set of finite automaton that includes a Stackelberg type. The method of proof is as follows: first we show that if the set of commitment types other than the Stackelberg type is small enough (i.e., $\mu(\Omega_-)$ is small) and the discount factor is sufficiently close to one, then player 1 receives payoff close to his highest individually rational payoff in any equilibrium (Lemma 1). Next, we show that player 1 can successfully manipulate player 2's beliefs. In particular, player 1 can repeatedly play the Stackelberg action and thereby ensure that player 2's posterior belief that player 1 is a type in Ω_- is arbitrarily small (Lemma 7). The main theorem, Theorem 1, puts the two findings together to show that player 1 can ensure a payoff close to his highest individually rational payoff by first manipulating player 2's beliefs so that the posterior belief of Ω_- is sufficiently low and then obtaining the high payoff outlined in Lemma 1.

Lemma 1 (Payoff Bound for Player 1)**.** *Posit Assumption 1 and Assumption 2. For any* $\underline{z}$ *and* $\gamma > 0$*, there exists a* $\underline{\delta}$ *and* $\bar{\phi}$ *such that, for any* $\delta > \underline{\delta}$*, any* μ *with* $\mu(\Omega_s) \geq \underline{z}$ *and* $\mu(\Omega_-) < \bar{\phi}$ *and any equilibrium strategy profile* σ *for the repeated game* $\Gamma^\infty(\mu, \delta)$*,* $U_1(\sigma) > \bar{g}_1 - \gamma = 1 - \gamma$.

The discussion that follows presents the various definitions and intermediate lemmas required for the proof Lemma 1. The proof of this central lemma is presented after all the intermediate results are established. First some preliminaries: Let a_2^b denote a best reply to a_1^s. By Assumption 2, $g_1(a_1^s, a_2^b) = \bar{g}_1$. Normalize payoffs, without loss of generality, such that:

(*i*) $\bar{g}_1 = 1$,

(*ii*) $\min_{a_2 \in A_2} g_1(a_1^s, a_2) \geq 0$,

(*iii*) $g_2(a_1^s, a_2^b) = 0$,

(*iv*) For $l > 0$, $g_2(a_1^s, a_2) \leq -l$ for any a_2 that is not a best reply to a_1^s.

The main focus of analysis in the proof of Lemma 1 is player 2's resistance against a Stackelberg type. Resistance is the expectation of the normalized discounted sum of the number of non-best replies player 2 will play against a Stackelberg type, in a particular equilibrium. Formally, the definition is as follows:

Definition 2 (Resistance)**.** *For action profile* $a = (a_1, a_2)$, *let* $i(a) = 1$ *if* a_2 *is not a best reply to* a_1^s *and* $i(a) = 0$, *otherwise. Let* $i(\delta, h_{-t}) = (1-\delta)\sum_{k=t}^{\infty} \delta^{k-t} i(a_k)$. *Let* $\sigma = (\sigma_1(s), \sigma_2)$. *Player 2's continuation resistance,* $R(\delta, \sigma_2|h_t) = \mathbb{E}_\sigma[i(\delta, h_{-t})|h_t]$ *where the expectation is over all histories generated by* σ *given* h_t. *Also, let* $R(\delta, \sigma_2) = R(\delta, \sigma_2|h_0)$.

The payoff to player 1 of using a strategy that always plays the Stackelberg action is at least $1 - R(\delta, \sigma_2)$, by the definition of resistance and normalization (*i*) and (*ii*). This trivially implies the following lemma.

Lemma 2. *In any Bayes-Nash equilibrium* σ *of* $\Gamma^\infty(\mu, \delta)$, $U_1(\sigma) \geq 1 - R(\delta, \sigma_2)$ *and* $U_1(\sigma|h_t) \geq 1 - R(\delta, \sigma_2|h_t)$ *for any* h_t *that has positive probability under* σ.

The goal is to show that $R(\delta, \sigma_2)$ is bound by $C \max\{1-\delta, \phi\}$, for some constant C, in any equilibrium σ of $\Gamma^\infty(\mu, \delta)$ where $\mu_1(\Omega_s) = \underline{z}$ and $\frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} = \phi$. Thus, if $\max\{1-\delta, \phi\}$ is close to zero, then $R(\delta, \sigma_2)$ is close to zero and $U_1(\sigma)$ is close to 1, in any equilibrium σ of $\Gamma^\infty(\mu, \delta)$. The following definition introduces some reputation thresholds, denoted z_n for a given resistance level $K^n \max\{1-\delta, \phi\}$.

Definition 3 (Reputation Thresholds)**.** *Fix $\delta < 1$, $K > 1$ and $\phi \geq 0$. Let $\epsilon = \max\{1-\delta, \phi\}$. For each $n \geq 0$, let*

$$z_n = \sup\{z : \exists \textit{Bayes-Nash equilibrium } \sigma \textit{ of } \Gamma^{\infty}(\mu, \delta), \quad \textit{where } \mu_1(\Omega_s) = z \textit{ and } \frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi, \textit{ such that } R(\delta, \sigma_2) \geq K^n \epsilon\}. \tag{2}$$

Also, define q_n such that

$$\frac{z_n}{1 - q_n} = z_{n-1}. \tag{3}$$

In words, z_n is the highest reputation level of player 1 for which there exists an equilibrium of $\Gamma^{\infty}(\mu, \delta)$ in which player 2's resistance exceeds $K^n \epsilon$. The definition and $K^n > K^{n-1}$ implies that $z_n \leq z_{n-1}$. The q_n's are real numbers that link the reputation thresholds z_n. To interpret q_n, suppose that player 2 believes player 1 to be the Stackelberg type with probability z_n. Also, suppose that the total probability that any of player 1's types plays an action different than a_1^s at least once over the next M periods is q_n. Consequently, if player 1 plays the Stackelberg action a_1^s in each of the M periods, then the posterior probability that player 2 places on player 1 being the Stackelberg type is $\frac{z_n}{1-q_n}$.

The development that follows will establish that $q_n \geq \underline{q} > 0$ for all n such that $z_n \geq \underline{z}$, and for all δ and ϕ, that is, the q_n's are uniformly (in δ and ϕ) bounded away from zero. If $q_n \geq \underline{q}$, then starting from $z_0 \leq 1$, there exists a n^* such that $z_{n^*} \leq \underline{z}$. Since $z_{n^*} \leq \underline{z}$, if the initial reputation level is $\underline{z}$, then the maximal resistance of player 2 is at most $K^{n^*}\epsilon$, which is of the order of $\max\{1 - \delta, \phi\}$. The following lemma formalizes this discussion.

Lemma 3. *Suppose that $q_n \geq \underline{q}$ for all δ, ϕ and all n such that $z_n \geq \underline{z}$. There exists n^* such that if $\max\{1 - \delta, \phi\} < \frac{\gamma}{K^{n^*}}$, then $U_1(\sigma) > 1 - \gamma$ for all Bayes-Nash equilibria σ of $\Gamma^{\infty}(\mu, \delta)$ with $\mu_1(\Omega_s) \geq \underline{z}$ and $\frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi$.*

Proof. Let n^* be the smallest the integer such that $(1 - \underline{q})^{n^*} < \underline{z}$. Since $\underline{q} > 0$ such an integer exists. For all δ and ϕ such that $z_0 \geq \underline{z}$, $z_{n^*} < \underline{z}$. Consequently, by Definition 3,

$R(\delta, \sigma_2) < K^{n^*}\epsilon$ in any equilibrium σ of $\Gamma^\infty(\mu, \delta)$ with $\mu_1(\Omega_s) \geq \underline{z}$ and $\mu_1(\Omega_-) = \phi$. For any δ and ϕ where $z_0 < \underline{z}$, by Definition 3, $R(\delta, \sigma_2) < \epsilon < K^{n^*}\epsilon$ in any equilibrium σ of $\Gamma^\infty(\mu, \delta)$ with $\mu_1(\Omega_s) = \underline{z}$ and $\mu_1(\Omega_-) = \phi$. Consequently, by Lemma 2, $U_1(\sigma) > 1 - K^{n^*}\epsilon$. So, if $\epsilon = \max\{1 - \delta, \phi\} < \frac{\gamma}{K^{n^*}}$, then $U_1(\sigma) > 1 - \gamma$. □

In order to show that $q_n \geq \underline{q} > 0$ lower and upper bounds are established for player 2's payoffs. The argument hinges on the tension between player 2's magnitude of resistance and the speed at which player 1 builds a reputation. If player 2 resists the Stackelberg type of player 1, then player 2 must be doing so in anticipation that player 1 deviates from the Stackelberg action. Otherwise player 2 could do better by best responding to the Stackelberg action. The more player 2 resists player 1, the more player 2 must be expecting player 1 to deviate from the Stackelberg action. However, if player 1 is expected to deviate from the Stackelberg action with high probability, then the normal type of player 1 can build a reputation rapidly by imitating the Stackelberg type.

The upper bound for player 2's payoff is calculated for a reputation level z close to the reputation threshold z_n in an equilibrium where player 2's resistance is approximately equal to the maximal resistance possible given the reputation level. The following formally defines maximal resistance for player 2.[5]

Definition 4 (Maximal Resistance). *For any $\xi > 0$, let $z_\xi = z_n - \xi$ and*

$$K_\xi = \sup\{k : \exists \text{ Bayes-Nash equilibrium } \sigma \text{ of } \Gamma^\infty(\mu, \delta), \text{ where } \mu_1(\Omega_s) = z \text{ and } \frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi, \text{ such that } R(\delta, \sigma_2) \geq k\epsilon \text{ and } z_n \geq z \geq z_\xi\}. \tag{4}$$

Observe that by the definition of K_ξ, there exists $z_n \geq z \geq z_\xi$ and an equilibrium strategy profile σ such that $R(\delta, \sigma_2) \geq (K_\xi - \xi)\epsilon$. Also, by the definition of K_ξ and the definition of z_n, $K_\xi \geq K^n$. The definition of z_n and $K_\xi \geq K^n$ implies that for any $z_n \geq z \geq z_\xi$, $R(\delta, \sigma_2) \leq$

[5]This further definition is required since it is not guaranteed that when $\mu_1(\Omega_s) = z_n$, there exists an equilibrium where resistance equals $K^n\epsilon$. However, by the definition of the threshold z_n, for z close to z_n there exists an equilibrium where resistance is close to $K^n\epsilon$.

$K_\xi \epsilon$ in any Bayes-Nash equilibrium strategy profile σ of $\Gamma^\infty(\mu,\delta)$ where $\mu_1(\Omega_s)=z$. The following Lemma establishes an upper bound on Player 2's payoff in any equilibrium where the resistance is at least $(K_\xi-\xi)\epsilon$.

Lemma 4. *Posit Assumption 2 for player 1. Pick any $z_n \geq z \geq z_\xi$ and Bayes-Nash equilibrium σ of $\Gamma^\infty(\mu,\delta)$ with $\mu_1(\Omega_s)=z$ and $\frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi$, such that $R(\delta,\sigma_2) \geq (K_\xi-\xi)\epsilon$. For the chosen equilibrium profile σ,*

$$U_2(\sigma) \leq \rho\epsilon(q_\xi K_\xi + (q_n+q_\xi)K^n + K^{n-1} + \frac{2(1-\delta)}{\epsilon}) - z(K_\xi-\xi)\epsilon l + \phi M. \tag{UB}$$

Proof. Pick $z_n \geq z \geq z_\xi$ and fix a Bayes-Nash equilibrium σ of the game $\Gamma^\infty(\mu,\delta)$ with $\mu(\Omega_s)=z$ and $\frac{\mu(\Omega_-)}{\mu(\Omega_s)} \leq \phi$, such that $R(\delta,\sigma_2) \geq (K_\xi-\xi)\epsilon$. For any history h, let $\Pr_{\sigma_1}\{\exists t \leq T : (a_1)_t \neq a_1^s | h\}$ denote the total probability that player 1 deviates from the Stackelberg action a_1^s in some period $t \leq T$ given that each type ω of player 1 uses strategy $\sigma_1(\omega)$, the outcomes of player 2's actions follow history h in periods 0 through $T-1$, and the measure over types is μ. Define integer valued random variable τ as follows

$$\tau(h) = \min\{T : \Pr_{\sigma_1}\{\exists t \leq T : (a_1)_t \neq a_1^s | h\} > q_\xi\} \tag{5}$$

Let $\tau(h)=\infty$, if the set that is minimized is empty. The random variable τ is a stopping time since it is H_t measurable and integer valued. In words, $\tau(h)$ is the first period T such that, the total probability with which player 1 is expected to take an action different than the Stackelberg action a_1^s at least once, in any period $t \leq T$, exceeds q_ξ. By definition, for any $T' < \tau(h)$, $\Pr_{\sigma_1}\{\exists t \leq T' : (a_1)_t \neq a_1^s | h\} \leq q_\xi$. Also, let

$$\tau_n(h) = \min\{T : \Pr_{\sigma_1}\{\exists t \leq T : (a_1)_t \neq a_1^s | h\} > q_n + q_\xi\}. \tag{6}$$

Let $\tau_n(h)=\infty$, if the set that is minimized is empty. The definition implies that $\tau_n(h) \geq \tau(h)$. Let $t(h)$ denote the first period player 1 deviates from a_1^s in history h and let $t(h)=\infty$ if player 1 never deviates from a_1^s in h. Assume that if $\tau(h)=\infty$, then $\tau(h)=\tau_n(h)$ and if $t(h)=\infty$, then $\tau_n(h) \leq t(h)$. Let event $E_{[0,\tau)} = \{h : t(h) < \tau(h)\} \cap \{\omega_N\}$, that is, the event that player 1 is the normal type and deviates from the Stackelberg action before random time τ. The definitions imply that $\Pr_\sigma(E_{[0,\tau)}) \le q_\xi$, i.e., the probability that player 1's normal type deviates from a_1^s in history h before random time τ is bounded above by q_ξ. Similarly, let $E_{[\tau,\tau_n)} = \{h : \tau(h) \le t(h) < \tau_n(h)\} \cap \{\omega_N\}$ and $E_{[\tau_n,\infty]} = \{h : \tau_n(h) \le t(h)\} \cap \{\omega_N\}$. Clearly, $\Pr_\sigma(E_{[\tau,\tau_n)}) \le q_\xi + q_n$ and $\Pr_\sigma(E_{[\tau_n,\infty]}) \le 1$. Also, observe that the sets $\{h : t(h) < \tau(h)\}$, $\{h : \tau(h) \le t(h) < \tau_n(h)\}$, and $\{h : \tau_n(h) \le t(h)\}$ partitions the set of all repeated game histories.

Suppose $\mu(\Omega_s) = z \ge z_\xi$ and $\frac{\mu(\Omega_-)}{\mu(\Omega_s)} \le \phi$. Trivially, $\frac{z_\xi}{1-q_n-q_\xi} > \frac{z_\xi}{(1-q_n)(1-q_\xi)} = z_{n-1}$. Consequently, if player 1 has played only a_1^s in partial history h_t of history $h = (h_t, h_{-t})$, then for any t, $\mu(\Omega_s|h_t) \ge z$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \le \phi$; for any $t \ge \tau(h)$, $\mu(\Omega_s|h_t) \ge z_n$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \le \phi$; for $t \ge \tau_n(h)$, $\mu(\Omega_s|h_t) \ge z_{n-1}$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \le \phi$.

Suppose $h \in E_{[0,\tau)}$. For any such history h player 1 has always played an action compatible with a_s^1 in partial history h_t. Consequently, $\mu_1(\Omega_s|h_t) \ge z$, $\frac{\mu_1(\Omega_-|h_t)}{\mu_1(\Omega_s|h_t)} \le \phi$ and hence player 2's resistance is at most $K_\xi\epsilon$, by Definition 4. So, by Lemma 2, $U_1(\sigma|h_t) \ge 1 - K_\xi\epsilon$. Note that, $(U_1(\sigma|h_t), U_2(\sigma|h_t)) \in F$, so $U_2(\sigma|h_t) \le \rho(1 - U_1(\sigma|h_t))$. Consequently,

$$U_2(\sigma|h_t, E_{[0,\tau)}) \le \rho K_\xi \epsilon. \tag{7}$$

Suppose $h \in E_{[\tau,\tau_n)}$. Suppose $t(h) = \tau(h)$. If in h player 1 has not deviated from the Stackelberg action until time $\tau(h)$, then player 1's equilibrium continuation payoff $U_1(\sigma|h_t)$ must be at least as large as $\delta(1 - K^n\epsilon)$. This is because player 1 can play a_1^s in period $\tau(h)$, get zero for the period by normalization (ii), increase his reputation to at least z_n and thereby guarantee a continuation payoff of at least $1 - K^n\epsilon$. If $t(h) > \tau(h)$, then $U_1(\sigma|h_t) \ge 1 - K^n\epsilon \ge \delta(1 - K^n\epsilon)$. This implies that player 2's continuation payoff is at most $\rho(1 - \delta(1 - K^n\epsilon)$. So, if player 2 is facing the normal type of player 1, and $h \in E_{[\tau,\tau_n)}$, then player 2's continuation payoff

$$U_2(\sigma|h_t, E_{[\tau,\tau_n)}) \le \rho(K^n\epsilon + (1-\delta)). \tag{8}$$

Suppose $h \in E_{[\tau_n,\infty]}$. If player 1 has not deviated from the Stackelberg action until period $\tau_n(h)$. In period $\tau_n(h)$, if player 1 plays action a_1^s, then his reputation will exceed z_{n-1} in the next period. Also, in any period after $\tau_n(h)$ his reputation exceeds z_{n-1}. Consequently, by the same reasoning as the previous paragraph, if player 2 is facing the normal type of player 1, and $h \in E_{[\tau_n,\infty]}$, then player 2's continuation payoff

$$U_2(\sigma|h_t, E_{[\tau_n,\infty]}) \leq \rho(K^{n-1}\epsilon + (1-\delta)). \tag{9}$$

Player 2 can get at most M against any other commitment type and this happens with probability $\phi z \leq \phi$. Since player 2's resistance is $(K_\xi - \xi)\epsilon$ in the equilibrium under consideration, she loses $(K_\xi - \xi)\epsilon l$ against the Stackelberg type, and this happens with probability z. The probability that player 1 is a normal type and takes action $a_1 \neq a_1^s$ for the first time in any period $t < \tau(h)$, i.e., $h \in E_{[0,\tau)}$, is at most q_ξ; and an upper-bound on player 2's continuation payoff, conditional on this event, is given by Equation (7). The probability that player 1 is a normal type and takes action $a_1 \neq a_1^s$ for the first time in any period $\tau(h) \leq t < \tau_n(h)$, i.e., $h \in E_{[\tau,\tau_n]}$, is at most $q_\xi + q_1$; and an upper-bound on player 2's continuation payoff is given by Equation (8). Finally, the probability that player 1 is a normal type and takes action $a_1 \neq a_1^s$ for the first time in any period $\tau_n(h) \leq t$, i.e., $h \in E_{[\tau_n,\infty]}$, is at most 1; and an upper-bound on player 2's continuation payoff is given by Equation (9). Player 2's continuation payoffs against the normal type of player 1 only exclude periods where player 1's normal type plays action a_1^s and player 2's per-period payoff is at most zero, by normalization (iii), in these periods. Consequently,

$$U_2(\sigma) \leq q_\xi \rho K_\xi \epsilon + (q_\xi + q_n)\rho K^n \epsilon + \rho K^{n-1}\epsilon - z(K_\xi - \xi)\epsilon l + 2\rho(1-\delta) + \phi M$$

delivering the required inequality. □

Lemma 5. *Posit Assumption 1 and Assumption 2 for player 1. Suppose that* $z_n \geq z \geq z_\xi$ *and that* $\mu_1(\Omega_s) = z$ *and* $\frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi$. *In any Bayes-Nash equilibrium* σ *of* $\Gamma^\infty(\mu, \delta)$,

$$U_2(\sigma) \geq -\frac{\rho}{\underline{\pi}}\epsilon(q_\xi K_\xi + (q_\xi + q_n)K^n + K^{n-1} + \frac{2(1-\delta)}{\epsilon}) - \phi M. \tag{LB}$$

Proof. Fix a Bayes-Nash equilibrium σ of $\Gamma^\infty(\mu, \delta)$ where $z_n \geq z \geq z_\xi$, $\mu_1(\Omega_s) = z$ and $\frac{\mu_1(\Omega_-)}{\mu_1(\Omega_s)} \leq \phi$. If Γ is a strictly conflicting interest game for player 1, then $U_2(\sigma) \geq \hat{g}_2 = g_2(a_1^s, a_2^b) = 0$ in any Bayes-Nash equilibrium which exceeds the right-hand side of Equation (LB).

Posit Assumption 1 and assume that Γ satisfies Assumption 2 for player 1 but is not a strictly conflicting interest game for player 1. We calculate the payoff of player 2 if she deviates and uses the following alternative repeated game strategy σ_2^*. Suppose that player 2 always plays a_2^b, a pure action, if player 1 has played the Stackelberg action a_1^s in every prior period of the repeated game and plays according to the strategy σ if player 1 has deviated from the Stackelberg action a_1^s in any prior period of the repeated game. Using this strategy player 2 will receive payoff equal to zero in any period where player 1 plays a_1^s.

Let strategy profile $\sigma^* = (\sigma_1, \sigma_2^*)$. Suppose that the stopping times τ, τ_n, and the events $E_{[0,\tau)}$, $E_{[\tau,\tau_n)}$ and $E_{[\tau_n,\infty]}$ are defined as in Lemma 4. Again, as in Lemma 4, $\Pr_{\sigma^*}(E_{[0,\tau)}) \leq q_\xi$, $\Pr_{\sigma^*}(E_{[\tau,\tau_n)}) \leq q_\xi + q_n$ and $\Pr_{\sigma^*}(E_{[\tau_n,\infty]}) \leq 1$. Also, if player 1 has played only a_1^s in partial history h_t of history $h = (h_t, h_{-t})$, then for any t, $\mu(\Omega_s|h_t) \geq z$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \leq \phi$; for any $t \geq \tau(h)$, $\mu(\Omega_s|h_t) \geq z_n$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \leq \phi$; for $t \geq \tau_n(h)$, $\mu(\Omega_s|h_t) \geq z_{n-1}$ and $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \leq \phi$.

Suppose $h \in E_{[0,\tau)}$ and $t = t(h)$. For any such history h player 1 has always played an action compatible with a_s^1 in partial history h_t. Consequently, $\mu(\Omega_s|h_t) \geq z$, $\frac{\mu(\Omega_-|h_t)}{\mu(\Omega_s|h_t)} \leq \phi$ and hence player 2's resistance is at most $K_\xi \epsilon$, by Definition 4. So, by Lemma 2,

$$U_1(\sigma|h_t) = \sum_{y \in Y} \pi_y(\sigma_2(h_t)) U_1(\sigma|y, a_1, h_t) \geq 1 - K_\xi \epsilon$$

for any $a_1 \neq a_1^s$ and $a_1 \in supp(\sigma_1(h_t))$. Note that, $U_1(\sigma|y, a_1, h_t), U_2(\sigma|y, a_1, h_t)) \in F$, so

$$U_2(\sigma|y, a_1, h_t) \geq -\rho(1 - U_1(\sigma|y, a_1, h_t)).$$

by Assumption 2 and Equation (1). Taking the expectation with respect to a_1 using player 1's equilibrium strategy implies that $U_2(\sigma|y, h_t) \geq -\rho(1 - U_1(\sigma|y, h_t))$. But, because we are considering alternative strategy σ_2^* and player 1 has played a_1^s in all previous periods, player 2 plays a_2^b in period t – which is not necessarily equal to the action chosen by Player 2's equilibrium strategy. However, Assumption 1 implies that $\pi_y(a) \geq \underline{\pi}$ for any $a \in A_2$ and $y \in Y_2$. Consequently,

$$\begin{aligned} U_2(\sigma|h_t, E_{[0,\tau)}) &= \sum_{y \in Y} \pi_y(a_2^b) U_2(y, \sigma|h_t) \geq -\rho \sum_{y \in Y} \pi_y(a_2^b)(1 - U_1(\sigma|y, h_t)) \\ &\geq -\rho \sum_{y \in Y} \frac{\pi_y(\sigma_2(h_n))}{\underline{\pi}}(1 - U_1(\sigma|y, h_t)) \geq -\frac{\rho}{\underline{\pi}} K_\xi \epsilon \end{aligned}$$

The event $E_{[0,\tau)}$ occurs with probability at most q_ξ. Following a similar reasoning to the above paragraph and Lemma 4 implies that $U_2(\sigma|h_t, E_{[\tau,\tau_n)}) \geq -\frac{\rho}{\underline{\pi}}(K_n\epsilon + (1 - \delta))$ and the event $E_{[\tau,\tau_n)}$ occurs with probability at most $q_\xi + q_n$; also $U_2(\sigma|h_t, E_{[\tau_n,\infty]}) \geq -\frac{\rho}{\underline{\pi}}(K_{n-1}\epsilon + (1 - \delta))$ and the event $E_{[\tau_n,\infty]}$ occurs with probability at most 1. Player 2's continuation payoffs against the normal type of player 1 only exclude periods t where player 1's normal type plays action a_1^s in period t and in all periods prior to t. However, because under σ_2^* player 2 plays a_2^b in these periods, player 2's per-period payoff is zero, by normalization (iii), in these periods. Also, Player 2 can get at least $-M$ against any other commitment type with probability at most ϕ gets zero against the Stackelberg type with probability at most z. Consequently,

$$U_2(\sigma) \geq -\frac{\rho}{\underline{\pi}} K_\xi \epsilon q_\xi - \frac{\rho}{\underline{\pi}} K^n \epsilon (q_\xi + q_1) - \frac{\rho}{\underline{\pi}} K^{n-1}\epsilon - 2\frac{\rho}{\underline{\pi}}(1 - \delta) - \phi M$$

delivering the required inequality. $\square$

Lemma 6. *Let* $\underline{q} = \underline{z}(\frac{l\underline{\pi}}{2\rho} - \frac{3}{\underline{z}K} - \frac{M\underline{\pi}}{\underline{z}\rho K})$ *and pick* K *such that* $\underline{q} > 0$*. If* $z_n \geq \underline{z}$*, then* $q_n \geq \underline{q}$*, for all* δ *and* ϕ*.*

Proof. Combining the lower bound for $U_2(\sigma)$, given by Equation (LB) established in Lemma 5, and the upper bound for $U_2(\sigma)$, given by Equation (UB) established in Lemma 4, and simplifying by canceling ϵ delivers

$$z(K_\xi - \xi)\epsilon l \leq 2\frac{\rho}{\underline{\pi}}(q_\xi K_\xi + (q_\xi + q_n)K^n + K^{n-1} + \frac{2(1-\delta)}{\epsilon}) + \frac{2\phi M}{\epsilon}.$$

Taking $\xi \to 0$ implies that $z \to z_n$, $q_\xi \to 0$. Also, $K_\xi \geq K^n$ for each ξ implies that $\lim_{\xi\to 0}(K_\xi - \xi) = \lim_{\xi\to 0} K_\xi \geq K^n$. Consequently,

$$z_n K^n \epsilon l \leq 2\frac{\rho}{\underline{\pi}}(q_n K^n + K^{n-1} + \frac{2(1-\delta)}{\epsilon}) + \frac{2\phi M}{\epsilon}.$$

Rearranging, $q_n \geq \frac{z_n l\underline{\pi}}{2\rho} - \frac{1}{K} - \frac{2(1-\delta)}{\epsilon K^n} - \frac{\phi M\underline{\pi}}{\epsilon\rho K^n}$. Recall that $\epsilon = \max\{1-\delta, \phi\}$ and $z_n \geq \underline{z}$ so

$$q_n \geq \underline{z}(\frac{l\underline{\pi}}{2\rho} - \frac{3}{\underline{z}K} - \frac{M\underline{\pi}}{\underline{z}\rho K}) = \underline{q} > 0$$

delivering the required inequality. □

Given Lemma 6, Lemma 3 can be applied to complete the proof of Theorem 1.

Proof of Lemma 1. By Lemma 6 if $z_n \geq \underline{z}$, then, for all δ, ϕ, $q_n \geq \underline{q}$. Consequently, by Lemma 3, if $\max\{1-\delta, \phi\} < \frac{\gamma}{K^{n^*}}$, then $U_1(\sigma) > 1-\gamma$ for all Bayes-Nash equilibria σ of $\Gamma^\infty(\mu, \delta)$ with $\mu_1(\Omega_s) \geq \underline{z}$ and $\mu_1(\Omega_-) = \phi$ as required. □

The following proposition shows that player 1 can make the probability that player 2 places on player 1 being a commitment type other than the Stackelberg type arbitrarily small by playing the Stackelberg action for T periods. The main point of the proposition is showing that this T can be chosen independently from player 2's strategy. The proposition gives a uniform lower bound on the speed of learning for player 2.

Lemma 7 (Uniform Learning)**.** *Suppose player 1 has played only* a_1^s *in history* h_t*, and let* $\Omega_s(h_t) \supset \Omega_s$ *denote the set of types that behave identical to a Stackelberg type given* h_t *and* *let $\Omega_-(h_t) \subset \Omega_-$ denote the set of commitment types not in $\Omega_s(h_t)$. For any $\phi > 0$, there exists a T such that, $\Pr_{(\sigma_1(s),\sigma_2)}\{h : \frac{\mu(\Omega_-(h_T)|h_T)}{\mu(\Omega_s(h_T)|h_T)} < \phi\} > 1-\phi$, for any strategy σ_2 of player 2.*

Proof. Step 1. The set of histories is viewed as a stochastic process and $\Pr_{(\sigma_1(s),\sigma_2)}$ as the probability measure over the set of histories H generated by $(\sigma_1(s), \sigma_2)$. We show that for each finite subset $W \subset \Omega_-$ and any $\varepsilon > 0$, there exists a T such that, $\Pr_{(\sigma_1^s,\sigma_2)}\{h : \mu_1(W \cap \Omega_-(h_T)|h_T) < \varepsilon\} > 1-\varepsilon$, for any strategy σ_2 of player 2. Proving this is sufficient for the result since W can be picked such that $\mu_1(W)$ is arbitrarily close to $\mu_1(\Omega_-)$ and $\mu_1(\Omega_s(h_T)|h_T) \geq \mu_1(\Omega_s|h_T) > \mu_1(\Omega_s)$.

Step 2. Each ω is a finite automaton. Consider the stochastic process over states generated by σ^s. Since $(a_1)_t = a_1^s$ for all t, the transitions between states depend only on the realizations of y_2. In particular, probability to transition from q_1 to q_2 after history h_t is as follows:

$$p(q_1, q_2, \sigma_2(h_t)) = \sum_{\{y_2 \in Y_2 : \tau(y_2, a_1^s, q_1) = q_2\}} \pi_{y_2}(\sigma_2(h_t)) \geq \sum_{\{y_2 \in Y_2 : \tau(y_2, a_1^s, q_1) = q_2\}} \underline{\pi}$$

Observe $p(q_1, q_2, \sigma_2(h_t)) > 0$ if and only if $\sum_{\{y_2 \in Y_2 : \tau(y_2, a_1^s, q_1) = q_2\}} \underline{\pi} > 0$. Consequently, if $p(q_1, q_2, \sigma_2(h_t)) > 0$ for some σ_2 and h_t pair, then $p(q_1, q_2, \sigma_2'(h_k)) \geq \underline{\pi}$ for all σ_2' and h_k. This implies that the state space Q_ω of any ω can be (uniquely) partitioned into transitory states (Q_ω^0) and a collection of disjoint ergodic sets (Q_ω^j) such that $Q_\omega = \cup_{i=0}^M Q_\omega^i$. This partition is independent of σ_2' and h_k because if $p(q_1, q_2, \sigma_2(h_t)) > 0$ for some $\sigma_2(h_t)$, then $p(q_1, q_2, \sigma_2'(h_k)) \geq \underline{\pi}$ for all σ_2' and h_k.

Step 3. Let $E(T, K, \varepsilon)$ denote the set of histories such that for any $h \in E(T, K, \varepsilon)$, after initial history h_l, all $\omega \in W$ have entered an ergodic subset of states $Q_\omega^i \subset Q_\omega$ and all states $q \in Q_\omega^i$ have been visited at least K times by period $l + T$. Since there are only a finite number of sets that must be considered for $\omega \in W$ (i.e, $|W| \max_{\omega \in W} |Q_\omega|$ in total), for each K and each $\varepsilon > 0$ there exists a finite time T such that the set $E(T, K, \varepsilon)$ has measure at least $1 - \varepsilon$ under σ^s, i.e., $\Pr_{\sigma^s}\{E(T, K, \varepsilon)\} > 1 - \varepsilon$. As a consequence of the above step this time T can be picked independently from player 2's strategy σ_2. That is $\Pr_{\sigma^s}\{E(T,K,\varepsilon)\} > 1-\varepsilon$ for any σ_2.

Step 4. Let $p_t^\omega(h)$ denote the probability that a_1^s is played in period t after history h_t, conditional on being type ω. Also, let $L_t^\omega(h) = \frac{p_t^\omega(h)}{p_t^{\Omega_s}(h)} L_{t-1}^\omega(h)$ and $L_0^\omega(h) = \frac{\mu_0(\omega)}{\mu_0(\Omega_s)}$. By Fudenberg and Levine (1992) Lemma 4.1, $L_t^\omega(h) = \frac{\mu(\omega|h_t)}{\mu(\Omega_s|h_t)}$ and (L_t^ω, H_t) is a supermartingale, under σ^s. Observe $p_t^{\Omega_s}(h) = 1$ for $\sigma^s - a.e.$ history. Let $L^\omega(K,\varepsilon)$ denote the set of histories such that either $L_T^\omega(h) < \varepsilon$ or $|p_t^\omega(h) - 1| < \varepsilon$ in all but K periods for any $T > K$. Fudenberg and Levine (1992) Theorem 4.1 implies that there exists a K_ω independent of σ_2 such that $\Pr_{\sigma^s}\{L^\omega(K_\omega,\varepsilon)\} > 1-\varepsilon$.

Step 5. Let $\xi = \min_{q\in\{q\in\cup_{\omega\in W}Q_w : f_{a_s}^\omega(q)\neq 1\}}(1 - f_{a_s}^\omega(q))$. That is, $f_{a_s}^\omega(q)$ is the probability that type ω plays a_s in state q and ξ is the minimum of the set of numbers $1 - f_{a_s}^\omega(q)$ over the set q such that $f_{a_s}^\omega(q)$ is different than 1. The minimum is well defined since each ω is a finite automaton, W is a finite set and $\{q \in \cup_{\omega\in W} Q_w : f_{a_s}^\omega(q) \neq 1\} \neq \emptyset$ because $W \subset \Omega_-$.

Step 6. Pick $\frac{\varepsilon}{2|W|} < \xi$. Pick K such that $K > K_\omega$ and $\Pr_{\sigma^s}\{L^\omega(K_\omega, \frac{\varepsilon}{2|W|})\} > 1 - \frac{\varepsilon}{2|W|}$ for all $\omega \in W$. Pick T such that $\Pr_{\sigma^s}\{E(T,K,\frac{\varepsilon}{2})\} > 1 - \frac{\varepsilon}{2}$. Consequently,

$$\Pr_{\sigma^s}\{E(T,K,\frac{\varepsilon}{2}) \cap (\cap_{\omega\in W} L^\omega(K_\omega, \frac{\varepsilon}{2|W|}))\} > 1-\varepsilon.$$

By Step 3, for any $h \in E(T,K,\frac{\varepsilon}{2}) \cap (\cap_{\omega\in W} L^\omega(K_\omega, \frac{\varepsilon}{2|W|}))$ all ω are in an ergodic set $Q_\omega^*(h) = Q_\omega^{i(\omega)}(h)$ and all ergodic states $q \in \cup_\omega Q_\omega^*$ have been visited more than K times by time T. By Step 4, either $L_T^\omega(h) < \frac{\varepsilon}{2|W|}$ or $|p_t^\omega(h) - 1| < \frac{\varepsilon}{2|W|} < \xi$ all but K times. However, by the definition of ξ in Step 5, either $L_T^\omega(h) < \frac{\varepsilon}{2|W|}$ or for any ω with $L_T^\omega(h) > \frac{\varepsilon}{2|W|}$ for all $q \in Q_\omega^*(h)$, $f_{a_s}(q) = 1$. That is, either $L_T^\omega(h) < \frac{\varepsilon}{2|W|}$ or $\omega \in \Omega_s(h_T)$. So all $\omega \in W$ with $\mu(\omega|h_T) > \frac{\varepsilon}{2|W|}$ are in $\Omega_s(h_T)$. Hence $\mu(W \cap \Omega_-(h_T)|h_T) < \frac{\varepsilon}{2}$ for any h $\in E(T,K,\frac{\varepsilon}{2}) \cap (\cap_{\omega\in W} L^\omega(K_\omega, \frac{\varepsilon}{2|W|}))$ delivering the result. □

The main theorem given below combines the finding of Proposition 1 and 2 to prove that player 1 will receive his highest individually rational payoff by repeatedly playing the Stackelberg action. The proof proceeds as follows: Proposition 2 implies that player 1 can ensure that the measure of the set of commitment types other than the Stackelberg types is smaller than $\underline{\phi}$ by time T. If the measure of the set of commitment types other than the Stackelberg types is smaller than $\bar{\phi}$, then Proposition 1 implies that Player 1's payoff is close to the highest individually rational payoff.

Theorem 1. *Posit Assumption 1 and Assumption 2. For any $\underline{z}$ and $\gamma > 0$, there exists a δ^* such that, for any $\delta > \delta^*$, any μ with $\mu(\Omega_s) \geq \underline{z}$ and any equilibrium strategy profile σ for the repeated game $\Gamma^\infty(\mu, \delta)$, $U_1(\sigma) > \bar{g}_1 - \gamma = 1 - \gamma$.*

Proof. Pick ϕ such that $K^{n^*}\phi + \phi < \gamma$ where K and n^* are defined as in Lemma 1. By Lemma 7 there exists T such that $\frac{\mu_1(\Omega_-(h_T)|h_T)}{\mu_1(\Omega_s(h_T)|h_T)} < \phi$ with probability $1 - \phi$ if player 1 always play a_1^s in history h_T. Consequently, if player 1 always play a_1^s in history h_T, then by Lemma 1, $U_1(\sigma|h_T) > 1 - K^{n^*} \max\{1 - \delta, \phi\}$ with probability $1 - \phi$. Because always playing a_1^s is a strategy available to player 1,

$$U_1(\sigma) \geq \delta^T(1 - \phi)(1 - K^{n^*} \max\{1 - \delta, \phi\}) \geq \delta^T(1 - K^{n^*} \max\{1 - \delta, \phi\} - \phi).$$

So, if $\delta > \underline{\delta} = \max\{(\frac{1-\gamma}{1-K^{n^*}\phi-\phi})^{\frac{1}{T}}, 1 - \phi\}$, then $U_1(\sigma) > 1 - \gamma$. □

References

Abreu, D., P. Dutta, and L. Smith (1994): "The Folk Theorem for Repeated Games: A Neu Condition," *Econometrica*, 62(4), 939–948.

Abreu, D., and F. Gul (2000): "Bargaining and reputation," *Econometrica*, 68(3), 85–117.

Abreu, D., and D. Pearce (2007): "Bargaining, reputation, and equilibrium selection in repeated games with contracts," *Econometrica*, 75(3), 653–710.

Atakan, A., and M. Ekmekci (2008): "Reputation with Long Run Players," *Working Paper*.

Billingsley, P. (1995): *Probability and Measure*. John Wiley and Sons, New York, NY.

Celantani, M., D. Fudenberg, D. Levine, and W. Pesendorfer (1996): "Maintaining a Reputation Against a Long-Lived Opponent," *Econometrica*, 64, 691–704.

Chan, J. (2000): "On the non-existence of reputation effects in two-person infinitely repeated games," Working paper.

Cripps, M., E. Dekel, and W. Pesendorfer (2005): "Reputation with equal discounting in repeated games with strictly conflicting interests," *Journal of Economic Theory*, 121, 259–272.

Cripps, M., and J. Thomas (1997): "Reputation and perfection in repeated common interest games," *Games and Economic Behavior*, 18, 141–158.

Fudenberg, D., and D. Levine (1989): "Reputation and equilibrium selection in games with a patient player," *Econometrica*, 57, 759–778.

——— (1992): "Maintaining a reputation when strategies are imperfectly observed," *Review of Economic Studies*, 59, 561–579.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.

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

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

Myerson, R. (1991): *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA.

Rubinstein, A., and A. Wolinsky (1995): "Remarks on infinitely repeated extensive-form games," *Games and Economic Behavior*, 9, 110–115.

Schmidt, K. (1993a): "Commitment through incomplete information in a simple repeated bargaining game," *Journal of Economic Theory*, 60, 114–139.

——— (1993b): "Reputation and equilibrium characterization in repeated games with conflicting interests," *Econometrica*, 61, 325–351.

Sorin, S. (1995): "A note on repeated extensive games," *Games and Economic Behavior*, 9, 116–123.

Wen, Q. (2002): "A Folk Theorem for Repeated Sequential Games," *Review of Economic Studies*, 69, 493–512.

MEDS, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208

E-mail address: a-atakan@kellogg.northwestern.edu

E-mail address: m-ekmekci@kellogg.northwestern.edu