REPUTATION WITH LONG RUN PLAYERS

ALP E. ATAKAN AND MEHMET EKMEKCI

ABSTRACT. Previous work shows that reputation results may fail in repeated games with equally patient players. We restrict attention to extensive-form stage games of perfect information. A reputation result is provided for repeated games with two long-run players that have equal discount factors. If player 1 is a dynamic Stackelberg type with positive probability, then that player receives the highest payoff, that is part of an individually rational payoff profile, in any perfect equilibrium, as agents become patient.

Keywords: Repeated Games, Reputation, Equal Discount Factor, Long-run Players.

JEL Classification Numbers: C73, D83.

Date: First draft, March, 2008. This revision, June, 2009.

1. INTRODUCTION AND RELATED LITERATURE

We consider an infinitely repeated game where two equally patient agents play an extensive-form stage-game of perfect information. The stage games that we allow are restricted: either the stage game is a locally non-conflicting interest game; or it is a game of strictly conflicting interest. In a locally non-conflicting interest game player 2 (she) receives a payoff that strictly exceeds her minimax value in the profile where player 1 (he) receives his highest payoff.[1] We assume that player 1's type is private information and he is a "dynamic Stackelberg type" committed to playing a certain repeated game strategy with positive probability.[2] We show that player 1 receives his highest payoff that is consistent with the individual rationality of player 2, in any perfect equilibrium, as the common discount factor converges to one, and the probability of player 1 being any other commitment type converges to zero.

We thank Martin Cripps, Eddie Dekel and Christoph Kuzmics for helpful discussions; our co-editor Larry Samuelson and three referees for detailed comments; and Umberto Garfagnini for research assistance.

[1]See Assumption 1.

[2]The dynamic Stackelberg payoff for player 1 is the highest payoff that he can guarantee in the repeated game through public pre-commitment to a repeated game strategy (a dynamic Stackelberg strategy); and a dynamic Stackelberg type is a commitment type that plays according to such a strategy.

A reputation result was first established for infinitely repeated games by Fudenberg and Levine (1989, 1992). They showed that if a patient player 1 plays a stage game against a myopic opponent (or a sequence of short-run players) and if there is positive probability that player 1 is a type committed to playing the Stackelberg action in every period, then in any equilibrium of the repeated game player 1 gets at least his static Stackelberg payoff.[3] However, in a game with a non-myopic opponent, player 1 may achieve a payoff that exceeds his static Stackelberg payoff by using a dynamic strategy that rewards or punishes player 2 (see Celantani et al. (1996)). Conversely, fear of future punishment or expectation of future rewards can induce player 2 to not best respond to a Stackelberg action and thereby force player 1 below his static Stackelberg payoff. The non-myopic player 2 may fear punishment either from another commitment type (Schmidt (1993) or Celantani et al. (1996)) or from player 1's normal type following the revelation of rationality (Celantani et al. (1996) section 5 or Cripps and Thomas (1997)).[4] Nevertheless, reputation results have also been established for repeated games where player 1 faces a non-myopic opponent, but one who is sufficiently less patient than player 1, by applying the techniques of Fudenberg and Levine (1989, 1992) (Schmidt (1993), Celantani et al. (1996), Aoyagi (1996), or Evans and Thomas (1997)).

Reputation results are fragile in repeated games in which a simultaneous-move stage game is played by equally patient agents. In particular, a reputation result obtains only if the stage game is a strictly conflicting interest game (Cripps et al. (2005)), or if there is a strictly dominant action in the stage game (Chan (2000)). For other simultaneous-move games, such as the common interest game, a folk theorem by Cripps and Thomas (1997) shows that any individually rational and feasible payoff can be sustained in perfect equilibria of the repeated game, if the players are sufficiently patient. In the equilibria Cripps and Thomas (1997) construct, player 2 refrains from best responding to the Stackelberg action in a given period because she fears a punishment phase will start in the event that she best responds and player 1 reveals rationality simultaneously (i.e., she is "caught"). In this construction, player 1 is able to provide incentives for player 2 to not best respond to the Stackelberg action while revealing rationality with a small probability in a given

[3]The static Stackelberg payoff for player 1 is the highest payoff he can guarantee in the stage game through public pre-commitment to a stage game action (a Stackelberg action). See Fudenberg and Levine (1989) or Mailath and Samuelson (2006), page 465, for a formal definition.

[4]Player 1 reveals rationality if he chooses a move that would not be chosen by any commitment type.

period (i.e., building a reputation slowly). This is because, although the chance that player 2 gets "caught" for playing a best response is low, the continuation payoff for player 2, if she is "caught," is sufficiently bad.

Recent work on reputation has focused on simultaneous-move stage games. In contrast, we restrict attention to extensive-form stage games of perfect information. For the class of games we consider, without incomplete information, the folk theorem of Fudenberg and Maskin (1986) applies, under a full dimensionality condition (see Wen (2002)). Also, if the normal-form representation of the extensive-form stage game we consider is played simultaneously, then under incomplete information, a folk theorem applies for a subset of the class of games we consider (Cripps and Thomas (1997)). Consequently, our reputation result covers a significantly larger class of games than those covered by previous reputation results for equally patient agents. Our main contribution is a reputation result for locally non-conflicting interest stage-games.

In the environment that we consider the techniques of Fudenberg and Levine (1989, 1992), which are commonly used to establish reputation results, are not available. Instead we use novel methods, inspired by the bargaining literature (Myerson (1991), section 8.8), to establish our reputation result. Our result hinges on perfect information at the decision nodes where player 1 reveals rationality and establishes tight bounds on player 2's continuation payoffs at these nodes. We show that perfect information and subgame perfection preclude the possibility that player 1 punishes player 2 for best responding to the Stackelberg strategy while building a reputation slowly, by limiting the range of continuation payoffs available for player 2.

As noted previously by Schmidt (1993), Celantani et al. (1996) or Evans and Thomas (1997) if there is a positive probability that player 1 is a commitment type, other than the Stackelberg type, then player 1 may be unable to build a reputation. This is because a non-myopic opponent may be unwilling to experiment nodes of the repeated game that must be sampled in order for player 1 to build a reputation. Previous work has addressed this issue by assuming that types are learned due to exogenous noise (Celantani et al. (1996) or Aoyagi (1996)); by restricting the class of games (Schmidt (1993)); or by considering more complicated types (Evans and Thomas (1997)). In the environment we consider, the presence of commitment types, other than the Stackelberg type, can also hinder player 1 from building a reputation. Consequently, our main reputation result is stated for the case where the probability that player 1 is another type is small compared to the probability that he is the Stackelberg type. In section 4.2 we discuss how this restriction on the relative probability of other types can be relaxed if the other commitment types are uniformly learnable, or if there is exogenous noise as in Celantani et al. (1996) or Aoyagi (1996).

The paper proceeds as follows: section 2 describes the model, section 3 presents the main reputation result, section 4 provides a discussion of the results, and the appendix contains proofs not included in the main text.

2. The Model

We consider a repeated game $\Gamma^{\infty}(\delta)$ in which a stage game Γ is played by players 1 and 2 in periods $t \in \{0, 1, 2, ...\}$ and the players discount payoffs using a common discount factor $\delta \in [0, 1)$. *The stage game* Γ *is a two-player finite game of perfect information, that is, all information sets of* Γ *are singletons* **(perfect information)**.

For the stage game, D is the set of nodes (decision nodes and terminal nodes), d is a typical element of D, $Y \subset D$ is the set of terminal nodes and y is a typical element of Y. The payoff function of player i is $g_i : Y \to \mathbb{R}$. The finite set of pure stage game actions for player i is A_i and the set of mixed stage game actions is $\mathcal{A}_i$.[5] For any action profile $a = (a_1, a_2) \in A_1 \times A_2$ there is a unique terminal history $y(a) \in Y$ under the path of play induced by a. With a slight abuse of notation we let $g_i(a) = g_i(y(a))$ for any $a \in A_1 \times A_2$.

In the repeated game Γ^{∞} players have perfect recall and can observe past outcomes. $Y^t \times D$ is the set of period $t \geq 0$ public histories and $\{y^0, y^1, ..., y^{t-1}, d\}$ is a typical element. $H^t \equiv Y^t$ is the set of period $t \geq 0$ public histories of terminal nodes and $\{y^0, y^1, ..., y^{t-1}\}$ is a typical element.

Types and Strategies. Before time 0 nature selects player 1 as a type ω from a countable set of types Ω according to common-knowledge prior μ. Player 2 is known with certainty to be a normal type that maximizes expected discounted utility. Ω contains a normal type for player 1 that we denote N. Player 2's belief over player 1's types, $\mu : \bigcup_{t=0}^{\infty} Y^t \times D \to \Delta(\Omega)$, is a probability measure over Ω after each period t public history.

[5]An action $a_i \in A_i$ is a contingent plan that specifies a move from the set of feasible moves for player i at any decision node d where player i is called upon to move.

A behavior strategy for player i is a function $\sigma_i : \bigcup_{t=0}^{\infty} H^t \to \mathcal{A}_i$ and Σ_i is the set of all behavior strategies. A behavior strategy chooses a mixed stage game action given any period t public history of terminal nodes. Each type $\omega \in \Omega \setminus \{N\}$ is committed to playing a particular repeated game behavior strategy $\sigma_1(\omega)$. A strategy profile $\sigma = (\{\sigma_1(\omega)\}_{\omega\in\Omega}, \sigma_2)$ lists the behavior strategies of all the types of player 1 and player 2. For any period t public history h^t and $\sigma_i \in \Sigma_i$, $\sigma_i|_{h^t}$ is the continuation strategy induced by h^t. For $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$, $\Pr_{(\sigma_1,\sigma_2)}$ is the probability measure over the set of (infinite) public histories induced by (σ_1, σ_2).

Payoffs. A player's repeated game payoff is the normalized discounted sum of the stage game payoffs. For any infinite public history $h = \{y^0, y^1, ..., \}$, $u_i(h, \delta) = (1-\delta)\sum_{k=0}^{\infty} \delta^k g_i(y^k)$, and $u_i(h^{-t}, \delta) = (1-\delta)\sum_{k=t}^{\infty} \delta^{k-t} g_i(y^k)$ where $h^{-t} = \{y^t, y^{t+1}, ...\}$. Player 1 and player 2's expected continuation payoff, following a period t public history, under strategy profile σ, $U_1(\sigma, \delta|h^t) = U_1(\sigma_1(N), \sigma_2, \delta|h^t)$ and

$$U_2(\sigma, \delta|h^t) = \sum_{\omega\in\Omega} \mu(\omega|h^t) U_2(\sigma_1(\omega), \sigma_2, \delta|h^t),$$

where $U_i(\sigma_1(\omega), \sigma_2, \delta|h^t) = \mathbb{E}_{(\sigma_1(\omega),\sigma_2)}[u_i(h^{-t}, \delta)|h^t]$ is the expectation over continuation histories h^{-t} with respect to $\Pr_{(\sigma_1(\omega)|_{h^t}, \sigma_2|_{h^t})}$. Also, $U_i(\sigma, \delta) = U_i(\sigma, \delta|h^0)$.

The stage game. The minimax payoff for player i, $\hat{g}_i = \min_{\alpha_j} \max_{\alpha_i} g_i(\alpha_i, \alpha_j)$. For games that satisfy perfect information there exists $a_1^p \in A_1$ such that $g_2(a_1^p, a_2) \le \hat{g}_2$ for all $a_2 \in A_2$.[6] 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 $G = F \cap \{(g_1, g_2) : g_1 \ge \hat{g}_1, g_2 \ge \hat{g}_2\}$. Let $\bar{g}_1 = \max\{g_1 : (g_1, g_2) \in G\}$, and $M = \max\{\max\{|g_1|, |g_2|\} : (g_1, g_2) \in F\}$.

Assumption 1. *The stage game* Γ *satisfies either of the following*

(*i*) *(Locally Non-Conflicting Interest) For any* $g \in G$ *and* $g' \in G$*, if* $g_1 = g_1' = \bar{g}_1$*, then* $g_2 = g_2' > \hat{g}_2$*, or*

(*ii*) *(Strictly Conflicting Interest) There exists* $a_1 \in A_1$ *such that any best response to* a_1 *yields payoffs* $(\bar{g}_1, \hat{g}_2)$*. Also,* $g_2 = \hat{g}_2$ *for all* $(\bar{g}_1, g_2) \in G$*.*[7]

[6] Consider the zero-sum game where player 1's payoff is equal to $-g_2(a_1, a_2)$. The minimax of this game is $(-\hat{g}_2, \hat{g}_2)$ by definition. Perfect information and Zermelo's lemma imply that this game has a pure strategy Nash equilibrium $(a_1^p, a_2) \in A_1 \times A_2$. Because the game is a zero sum game $g_2(a_1^p, a_2) = \hat{g}_2$.

[7] See Cripps et al. (2005), or Mailath and Samuelson (2006), page 541.

Assumption 1 requires that the payoff profile where player 1 obtains $\bar{g}_1$ is unique (i.e., the game is *generic*). Items (i) and (ii) are mutually exclusive. Item (i) requires that the game have a common value component: in the payoff profile where player 1 receives his highest payoff player 2 receives a payoff that strictly exceeds her minimax value. In contrast, item (ii) requires that the action which is the best for player 1 is the worst for his opponent.

Assumption 1 implies that there exists an action profile $(a_1^s, a_2^b) \in A_1 \times A_2$ such that $g_1(a_1^s, a_2^b) = \bar{g}_1$. If Γ is a strictly conflicting interest game we further take a_2^b to denote a best response to a_1^s.[8] If Γ satisfies Assumption 1 (i), then there exists $\rho \geq 0$ such that

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

If Γ satisfies Assumption 1 (ii), then there exists $\rho \geq 0$ such that

$$g_2 - g_2(a_1^s, a_2^b) \leq \rho(\bar{g}_1 - g_1), \text{ for any } (g_1, g_2) \in F. \tag{2}$$

We normalize payoffs, without loss of generality, such that

$$\bar{g}_1 = 1;\ g_1(a_1, a_2) \geq 0 \text{ for all } a \in A; \text{ and } g_2(a_1^s, a_2^b) = 0. \tag{3}$$

Dynamic Stackelberg payoff, strategy and type. Let

$$U_1^s(\delta) = \sup_{\sigma_1 \in \Sigma_1} \inf_{\sigma_2 \in BR(\sigma_1, \delta)} U_1(\sigma_1, \sigma_2, \delta),$$

where $BR(\sigma_1, \delta)$ denotes the set of best responses of player 2 to the repeated game strategy σ_1 of player 1. Let $\sigma_1^s(\delta)$ denote a strategy that satisfies $\inf_{\sigma_2 \in BR(\sigma_1^s(\delta), \delta)} U_1(\sigma_1^s(\delta), \sigma_2, \delta) = U_1^s(\delta)$, if such a strategy exists. We call $U_1^s(\delta)$ the dynamic Stackelberg payoff and $\sigma_1^s(\delta)$ a dynamic Stackelberg strategy for player 1.[9] If Assumption 1 is satisfied by Γ, then $U_1^s(\delta) = 1$ and a dynamic Stackelberg strategy exists in the repeated game $\Gamma^\infty(\delta)$ for all δ that exceed a cutoff $\delta^* \in [0, 1)$. The dynamic Stackelberg payoff, which we define for the repeated game, may exceed the static Stackelberg payoff (for a definition of the static Stackelberg payoff see Fudenberg and Levine (1989) or Mailath and Samuelson (2006)).

[8]If Γ is a locally non-conflicting interest game a_2^b is not necessarily a best response to a_1^s.
[9]The terminology follows Aoyagi (1996) and Evans and Thomas (1997).

We focus on a particular Stackelberg type, denoted S, that plays a strategy $\sigma_1(S)$. The strategy $\sigma_1(S)$ has a profit phase and a punishment phase. In the profit phase the strategy plays a_1^s and in the punishment phase the strategy plays a_1^p. The strategy begins the game in the profit phase. The strategy remains in the profit phase in period t if it was in the profit phase in period $t-1$ and $g_1(y_{t-1}) = 1$. The strategy moves to the punishment phase in period t if it was in the profit phase in period $t-1$ and $g_1(y_{t-1}) \neq 1$. If the strategy moves to the punishment phase in period t, then it remains in the punishment phase for $n^p - 1$ periods and then moves to the profit phase. Intuitively, $\sigma_1(S)$ punishes player 2, by minimaxing her for the next $n^p - 1$ periods, if she does not allow player 1 to obtain a payoff of one. The number of punishment periods $n^p - 1$ is the smallest integer such that

$$g_2(a_1^s, a_2) + (n^p - 1)\hat{g}_2 < n^p g_2(a_1^s, a_2^b) = 0 \tag{4}$$

for any $a_2 \in A_2$ such that $g_1(a_1^s, a_2) < g_1(a_1^s, a_2^b) = 1$. Assumption 1 implies that $n^p \geq 1$ exists. The number of punishment periods is chosen to ensure that it is a best response for a sufficiently patient player 2 to play a_2^b in every period against $\sigma_1(S)$. That is, if $\sigma_2 \in BR(\sigma_1(S), \delta)$, then $U_1(\sigma_1(S), \sigma_2, \delta) = 1$, for sufficiently high δ. Consequently, $\sigma_1(S)$ is a dynamic Stackelberg strategy for sufficiently high δ.

If $n^p = 1$, then S is a simple Stackelberg type and the static Stackelberg payoff coincides with the dynamic Stackelberg payoff for any discount factor (see Figure 1).[10] If $n^p > 1$, then the dynamic Stackelberg payoff strictly exceeds the static Stackelberg payoff for a sufficiently high discount factor (see Figure 2).

In what follows we assume that Ω *contains the Stackelberg type* S. Let $\Omega_- = \Omega \backslash \{S, N\}$. In words, Ω_- is the set of types other than the Stackelberg type and the normal type.

Equilibrium and beliefs. The repeated game where the prior over Ω is μ and the discount factor is δ is denoted $\Gamma^\infty(\mu, \delta)$. The analysis in the paper focuses on the perfect Bayesian equilibria (PBE) 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.

[10]In our figures the first component of the payoff vector is player 1's payoff and the second is player 2's payoff.

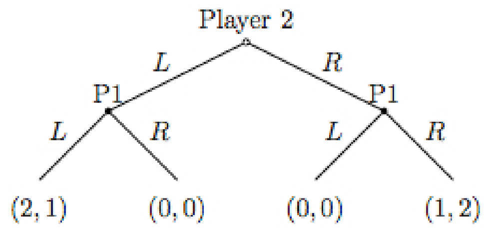

FIGURE 1. In this game S plays L regardless of Player 2's move and $n^p = 1$.

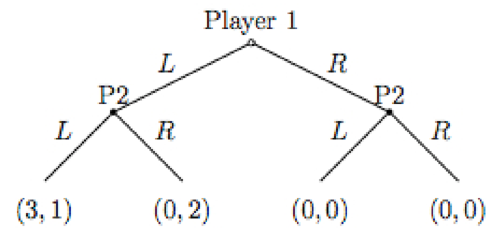

FIGURE 2. In this game S plays L in the profit phase and R in the two period punishment phase and $n^p = 3$.

If $\mu(S) > 0$, then belief $\mu(\cdot|h^t)$ is well defined after any period t public history where player 1 has played according to $\sigma_1(S)$.

3. The Main Reputation Result

Theorem 1 restricts attention to stage games of perfect information that satisfy Assumption 1 and considers a repeated game $\Gamma^\infty(\mu, \delta)$ where $\mu(S) > 0$. It demonstrates that a normal type for player 1 can secure a payoff arbitrarily close to one, his dynamic Stackelberg payoff, by mimicking the Stackelberg type, in any PBE of the repeated game, for a sufficiently high discount factor ($\delta > \underline{\delta}$) and for sufficiently low probability mass on other commitment types ($\mu(\Omega_-) < \bar{\phi}$).

Theorem 1. *Posit perfect information and Assumption 1. For any $\underline{z} > 0$ and $\gamma > 0$, there exists $\underline{\delta} < 1$ and $\bar{\phi} > 0$ such that, for any $\delta \in (\underline{\delta}, 1)$, any $\mu \in \Delta(\Omega)$ with $\mu(S) \geq \underline{z}$ and $\mu(\Omega_-) < \bar{\phi}$ and any PBE strategy profile σ of $\Gamma^\infty(\mu, \delta)$, $U_1(\sigma) > 1 - \gamma$.*

We begin by introducing the definitions used in the proof of Theorem 1. Let the resistance of strategy σ_2

$$r(\sigma_2, \delta) = 1 - U_1(\sigma_1(S), \sigma_2, \delta).$$

Below we define the maximal resistance function, $R(\mu, \delta)$, which is an upper-bound on how much player 2 can resist (or hurt) type S in any PBE of $\Gamma^\infty(\mu, \delta)$.

Definition 1 (Maximal resistance function)**.** *For any measure $\mu \in \Delta(\Omega)$ and $\delta \in [0, 1)$ let*

$$R(\mu, \delta) = \sup\{r(\sigma_2, \delta) : \sigma_2 \text{ is part of a PBE profile } \sigma \text{ of } \Gamma^\infty(\mu, \delta)\}.$$

We say that player 1 deviated from $\sigma_1(S)$ in the t^{th} period of a public history h if there exists a node d within period t where the move of player 1 differs from the move that strategy $\sigma_1(S)$ would have chosen at that node. For an integer T, $E_{[0,T]}$ denotes the event (set of infinite public histories) where player 1 deviates from $\sigma_1(S)$ for the first time in period t for some $0 \leq t \leq T$.

The following defines, $T(\sigma,\mu,q)$, the first period where the total probability of a deviation from $\sigma_1(S)$ by any of player 1's types exceeds q.

Definition 2 (Stopping time)**.** *For any strategy profile $\sigma = (\{\sigma_1(\omega)\}_{\omega\in\Omega}, \sigma_2)$ where σ_2 is a pure strategy, measure $\mu \in \Delta(\Omega)$ and $q \in [0,1]$ let*

$$T(\sigma,\mu,q) = \min\{T : \sum_{\omega\in\Omega} \mu(\omega) \Pr_{(\sigma_1(\omega),\sigma_2)} \left[E_{[0,T]}\right] > q\},$$

and let $T(\sigma,\mu,q) = \infty$ if the set is empty.[11]

Suppose that player 1's initial reputation level $\mu(S) = z$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$. If $h^{T(\sigma,\mu,q)+1}$ is a history consistent with $\sigma_1(S)$ and σ_2, i.e., player 1 has not deviated from $\sigma_1(S)$ in $h^{T(\sigma,\mu,q)+1}$, then Bayes' rule implies that

$$\mu(S|h^{T(\sigma,\mu,q)+1}) > \frac{z}{1-q} \text{ and } \frac{\mu(\Omega_-|h^{T(\sigma,\mu,q)+1})}{\mu(S|h^{T(\sigma,\mu,q)+1})} \leq \phi.$$

In this section we assume perfect information and Assumption 1. Additionally, we assume that $n^p = 1$ and $\mu(\Omega_-) = 0$ and we maintain Assumption 2 given below. Notice $n^p = 1$ implies that a_2^b is a best reply to a_1^s and normalization (3) implies that if $a_2 \notin BR(a_1^s)$, then $g_2(a_1^s, a_2) < 0$. Also, because $\mu(\Omega_-) = 0$, we let $z \in [0,1]$ represent the measure μ. One should understand this to mean $\mu(S) = z$ and $\mu(N) = 1 - z$.

Assumption 2. *$R(z,\delta)$ is a non-increasing function of z for each $\delta \in [0,1)$.*

The additional assumptions ($n^p = 1$, $\mu(\Omega_-) = 0$ and Assumption 2) are for expositional convenience only and allow us to present the main ingredients of Theorem 1's proof. The complete argument for Theorem 1 as well as a brief description of the argument is in the Appendix.

The payoff to player 1 of using the Stackelberg strategy is at least $1 - R(z,\delta)$. We argue that $\lim_{\delta\to 1} R(z,\delta) = 0$, for any $z > 0$. We begin by bounding continuation payoffs, at any decision node d where player 1 deviates from $\sigma_1(S)$, in terms of $R(z,\delta)$.

[11]If σ_2 is not a pure strategy, then $T(\sigma,\mu,q)$ is a random stopping time.

Lemma 1. *Pick any PBE σ of $\Gamma^{\infty}(z,\delta)$, period t public history $h^t = (h^{t-1}, d_0)$, and suppose player 1 is to deviate from $\sigma_1(S)$ at node d_0 with positive probability given h^t. Let $h = (h^{t-1}, d)$ be any public history that is reached with positive probability immediately following the deviation; let $h' = (h^{t-1}, d')$ be the public history that is reached immediately if $\sigma_1(S)$ is used at d_0 instead of deviating; and let $z' = \mu(S|h')$; then $U_1(\sigma,\delta|h) \geq 1 - R(z',\delta) - 2M(1-\delta)$. Consequently,*

$$|U_2(\sigma_1(N), \sigma_2, \delta|h)| \leq \rho(R(z',\delta) + 2M(1-\delta)), \text{ if } \Gamma \text{ satisfies Ass. 1 (i), and}$$

$$U_2(\sigma_1(N), \sigma_2, \delta|h) \leq \rho(R(z',\delta) + 2M(1-\delta)), \text{ if } \Gamma \text{ satisfies Ass. 1 (ii).}$$

Proof. If player 1 plays according to $\sigma_1(S)$ at d_0, then his reputation level $\mu(S|h') = z'$. After history h', player 1 can guarantee continuation payoff equal to $1 - R(z',\delta)$ by using $\sigma_1(S)$. Also, player 1 can get at worst zero in period t by playing according to $\sigma_1(S)$. Consequently, he can guarantee $\delta(1 - R(z',\delta)) \geq 1 - R(z',\delta) - M(1-\delta)$. If instead player 1 deviates from $\sigma_1(S)$, then he can get at most $M(1-\delta)$ for the period and $U_1(\sigma,\delta|h)$ as the continuation payoff. Since playing according to $\sigma_1(S)$ is a feasible alternative for player 1, $M(1-\delta)+U_1(\sigma,\delta|h) \geq 1-R(z',\delta)-M(1-\delta)$, or, $U_1(\sigma,\delta|h) \geq 1 - R(z',\delta) - 2M(1-\delta)$. The bounds on player 2's payoff follow from equations (1), (2) and $(U_1(\sigma,\delta|h), U_2(\sigma_1(N),\sigma_2,\delta|h)) \in F$. □

In order to provide some intuition for Lemma 1, we contrast the implications of Lemma 1 in a particular equilibrium of a simultaneous-move game depicted in Figure 3, with the implications of the lemma for a game of perfect information depicted in Figure 4.[12]

Take the game in Figure 3 and suppose that player 1 is a Stackelberg type that plays U in every period, with probability z. We construct a PBE for this game, following Cripps and Thomas (1997), where the players' payoffs are low, if z is close to zero and δ is close to one. Suppose player 2 plays R and player 1 uses a mixed strategy that plays D with small probability for the first K periods. After the first K periods (L,U) is played forever. In this construction $U_1(\sigma) = U_2(\sigma) = \delta^K$. Also, both players' continuation payoff, after (R,D) or (R,U), is equal to δ^{K-t} in any period $t \in \{0,..,K-1\}$. To ensure that player 2 has an incentive to play R, she is punished in the event that she plays

[12]In Figure 4 if $\epsilon = 0$, then in the complete information repeated game, i.e., without commitment types, the unique perfect equilibrium involves both agents receiving a payoff equal to one (see Rubinstein and Wolinsky (1995)). In order for an interesting reputation result we take $\epsilon > 0$ which ensures that G has a non-empty interior and implies that any payoff profile in the interior of G is sustained as a perfect equilibrium of the complete information repeated game for sufficiently patient players.

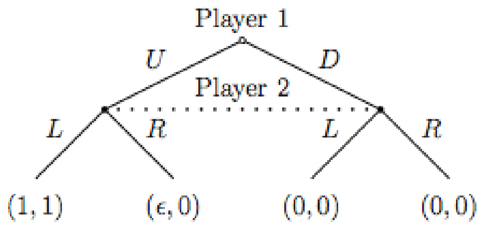

FIGURE 3. A simultaneous move common interest game ($\epsilon < 1$) that does not satisfy perfect information.

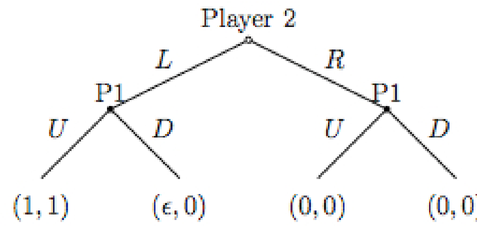

FIGURE 4. A sequential-move common interest game ($0 < \epsilon < 1$) that satisfies perfect information.

L and player 1 plays D (thus revealing rationality). Punishment entails a continuation payoff for player 2 that is close to zero.[13] Player 1 is willing to mix between U and D in the first K periods since player 2 only plays R on the equilibrium path.

In this construction, by choosing player 2's continuation payoff close to zero at nodes (L, D), she can be deterred from playing L even if player 1 reveals rationality with a small probability in each period. However, if the probability that player 1 reveals rationality is small in each period, then it takes many periods for player 1 to build a reputation and K can be chosen large to ensure low payoffs for both players. This argument hinges on choosing low continuation payoffs for player 2 after terminal node (L, D) during the first K periods. In the first K periods when player 1 makes his move he expects player 2 to play L with probability zero. Consequently, *the terminal node* (L, D) *is reached with probability zero and Lemma 1 puts no restrictions on payoffs at* (L, D).

Suppose instead that player 1 moves after observing player 2's choice as in Figure 4. Take any public history $h^t = (h^{t-1}, L)$ where player 1's reputation is $z > 0$ and player 1 has just observed L. Suppose that player 1's equilibrium strategy entails revealing rationality by playing D with probability $q > 0$ given h^t. After observing L, player 1 can always play U (not reveal rationality), receive $1-\delta$ for the period, and ensure a continuation payoff of at least $\delta(1-R(\frac{z}{1-q}, \delta))$. Consequently, player 1's payoff if he plays D, that is, his payoff at node (L, D) must be at least $1 - \delta R(\frac{z}{1-q}, \delta)$. This argument is a restatement of the lower-bound that Lemma 1 puts on player 1's payoff at node (h^t, L, D). In contrast to the simultaneous-move game, even if player 1 expects player 2 to play L with probability zero on the equilibrium path, but nevertheless observes L, player

[13]After (L, D) or (R, D) we are in a repeated game of complete information and any payoff in $[0, 1]$ can be supported.

1 realizes that the game has deviated from equilibrium and the bound on his payoff at (L, D) follows from applying perfection to this subgame. If player 1's payoff is at least $1 - \delta R(\frac{z}{1-q}, \delta)$ at node (L, D), then player 2's payoff at node (L, D) must be at least $1 - \frac{\delta}{1-\epsilon} R(\frac{z}{1-q}, \delta)$, by equation (1) and because $\rho = \frac{1}{1-\epsilon}$ in this game. Notice, if $R(\frac{z}{1-q}, \delta)$ is small to begin with, Lemma 1 greatly restricts the payoffs for player 2 after (L, D) in Figure 4. So, in contrast to the simultaneous-move version of the game, Lemma 1 bounds the amount of punishment that player 2 can expect after choosing L.

In what follows we focus on an equilibrium where player 2 resists the Stackelberg type by approximately $R(z, \delta)$. We compare player 2's payoff in this equilibrium with her payoff if she uses an alternative strategy that best responds to the Stackelberg strategy until player 1 reveals rationality. Resisting is costly for player 2 since there is a positive probability that she actually faces the Stackelberg type. The alternative strategy allows player 2 to avoid this cost. However, player 2 may be resisting the Stackelberg type because she expects a reward in the event that she sticks with equilibrium play and player 1 reveals rationality; or because she fears punishment in the event that she best responds to the Stackelberg strategy and player 1 reveals rationality. Lemma 2 below gives an upper-bound on player 2's payoff if she sticks to the equilibrium strategy. The lemma ties player 2's expected reward to the maximal resistance function using Lemma 1. Also, the lemma takes into account that player 2 bears a cost against the Stackelberg type. Lemma 3 gives a lower-bound on player 2's payoff if she uses the alternative strategy. Similarly this lemma ties player 2's expected punishment to the maximal resistance function using Lemma 1.

Lemma 2 (Upper-bound)**.** *Suppose* $0 \le z < z' \le 1$. *Let* σ *denote a PBE of* $\Gamma^\infty(z, \delta)$ *where player 2's resistance is at least* $R(z, \delta) - \epsilon$ *for* $\epsilon > 0$. *Then,*

$$U_2(\sigma, \delta) \le \rho(qR(z, \delta) + R(z', \delta) + 4M(1-\delta)) - zl(R(z, \delta) - \epsilon) \tag{5}$$

where $q = 1 - z/z'$ *and* $l > 0$ *is such that if* $a_2 \notin BR(a_1^s)$, *then* $g_2(a_1^s, a_2) < -l$.

Proof. For any $\epsilon > 0$ we can find such a PBE. Let σ_2^* denote a pure strategy in the support of σ_2 whose resistance is at least $R(z(1-q), \delta) - \epsilon$, let profile $\sigma^* = (\sigma_1(N), \sigma_1(S), \sigma_2^*)$ and let $T = T(\sigma^*, z, q)$. Since the resistance of σ_2 is at least $R(z, \delta) - \epsilon$, there must be a pure strategy in the support of σ_2 that has resistance of at least $R(z,\delta)-\epsilon$. T is the first period that the total probability that player 1 deviates from $\sigma_1(S)$ exceeds q, if the initial probability that player 1 is type S equals z. If player 1 does not deviate from $\sigma_1(S)$ until time $T+1$, then the posterior probability that player 1 is the Stackelberg type exceeds z'. We look at three events. The event that player 1 deviates from $\sigma_1(S)$ for the first time at $t<T$ (event $E_{[0,T)}$), the event that player 1 deviates from $\sigma_1(S)$ for the first time at $t \geq T$ (event $E_{[T,\infty)}$), and the event that player 1 is the Stackelberg type.

Player 2's payoff until player 1 deviates from $\sigma_1(S)$ is at most zero by normalization (3). Player 2's continuation payoff after player 1 deviates from $\sigma_1(S)$ is at most $\rho(R(z,\delta)+2M(1-\delta))$ if the deviation occurs at $t<T$; and is at most $\rho(R(z',\delta)+2M(1-\delta))$ if the deviation occurs at $t \geq T$, by Lemma 1. Consequently, $U_2(\sigma^*,\delta|E_{[0,T)}) \leq \rho(R(z,\delta)+2M(1-\delta))$ and $U_2(\sigma^*,\delta|E_{[T,\infty)}) \leq \rho(R(z',\delta)+2M(1-\delta))$. The expected, discounted number of periods where player 1 gets a payoff less than one is at least $R(z,\delta)-\epsilon$. So, the expected discounted number of periods where player 2 does not best respond to a_1^s is at least $R(z,\delta)-\epsilon$. Consequently, $U_2(\sigma_1(S),\sigma_2^*,\delta) \leq -l(R(z,\delta)-\epsilon)$ (see also Lemma 5). The probability of event $E_{[0,T)}$ is at most q, probability of event $E_{[T,\infty)}$ is at most one, and the probability of S is equal to z. Also, $U_2(\sigma,\delta)=U_2(\sigma^*,\delta)$. So,

$$U_2(\sigma,\delta) \leq \rho(qR(z,\delta)+R(z',\delta)+4M(1-\delta)) - zl(R(z,\delta)-\epsilon).$$

□

In Lemma 2, player 2's payoff is bounded along the equilibrium path. Consequently, in this lemma the perfect information assumption is not required. Consider again the equilibrium described for the simultaneous-move game depicted in Figure 3. Lemma 2 uses Lemma 1 to bound players 2's payoff at terminal nodes (R,D), which are reached along the equilibrium path with positive probability, during the first K periods.

Lemma 3 (Lower-bound)**.** *Suppose* $0 \leq z < z' \leq 1$*. In any PBE* σ *of* $\Gamma^\infty(z,\delta)$

$$U_2(\sigma,\delta) \geq -\rho(qR(z,\delta)+R(z',\delta)+4M(1-\delta)) \tag{6}$$

where $q = 1 - z/z'$*.*

Proof. Pick any PBE σ of $\Gamma^{\infty}(z,\delta)$ and suppose Assumption 1 (i) holds. Let σ_2^* denote a strategy that moves according to a_2^b after any period k public history h^k, if there is no deviation from $\sigma_1(S)$ in h^k, and coincides with PBE strategy σ_2 if player 1 has deviated from $\sigma_1(S)$ in h^k. Let strategy profile $\sigma^* = (\sigma_1(N), \sigma_1(S), \sigma_2^*)$ and let $T = T(\sigma^*, z, q)$. We again look at the events $E_{[0,T)}$, $E_{[T,\infty)}$, and the event that player 1 is type S. Player 2's payoff until player 1 deviates from $\sigma_1(S)$ is zero by normalization (3). Consequently, $U_2(\sigma^*, \delta|E_{[0,T)}) \geq -\rho(R(z,\delta) + 2M(1-\delta))$ and $U_2(\sigma^*, \delta|E_{[T,\infty)}) \geq -\rho(R(z',\delta) + 2M(1-\delta))$. Also, $U_2(\sigma^*|S) = 0$ by the definition of σ_2^*. So,

$$U_2(\sigma,\delta) \geq U_2(\sigma^*,\delta) \geq -\rho(qR(z,\delta) + R(z',\delta) + 4M(1-\delta)).$$

If Assumption 1 (ii) holds, then $U_2(\sigma,\delta) \geq \hat{g}_2 = 0 \geq -\rho(qR(z,\delta) + R(z',\delta) + 4M(1-\delta))$. □

In contrast to Lemma 2, perfect information is essential for Lemma 3. To see this, again consider the games depicted in Figures 3 and 4. In Lemma 3 we consider a strategy for player 2 that plays L until player 1 deviates from U and we give a lower-bound for player 2's payoff after (L,D). As we discussed previously, Lemma 1 only provides a lower-bound on player 2's payoff after (L,D) in the case of perfect information. Lemma 1 puts no restrictions on payoffs after (L,D) in the PBE we construct for the simultaneous-move game.

Below we use the fact that the upper-bound provided in Lemma 2 must exceed the lower-bound given in Lemma 3 to obtain a functional inequality that relates maximal resistance at any two reputation levels. We then use this functional inequality to complete our proof.

Lemma 4 (Functional Inequality)**.** *For any* $z \in [\underline{z}, 1]$ *and* $z < z' \leq 1$

$$R(z,\delta)(\underline{z}l - q(l+2\rho)) \leq 2\rho R(z',\delta) + 8\rho M(1-\delta) \tag{7}$$

where $q = 1 - z/z'$.

Proof. Combining (5) and (6), taking $\epsilon \to 0$, and simplifying delivers inequality (7). □

Proof of Theorem 1 under perfect information, Assumption 1, $n^p = 1$, $\mu(\Omega_-) = 0$ and Assumption 2. Let $\underline{q} = \frac{\underline{z}l}{2(l+2\rho)}$. For any $z \in [\underline{z}, 1]$, $z < z' \leq 1$ such that $1 - z/z' = q \leq \underline{q}$ inequality (7) implies

$$R(z,\delta)(\underline{z}l - \underline{q}(l+2\rho)) \leq 2\rho R(z',\delta) + 8\rho M(1-\delta)$$

$$R(z,\delta) \leq \frac{4\rho}{\underline{z}l}(R(z',\delta) + 4M(1-\delta)) \text{ (substituting } \frac{\underline{z}l}{2(l+2\rho)} \text{ for } \underline{q})$$

$$R(z'(1-q),\delta) \leq \frac{4\rho}{\underline{z}l}(R(z',\delta) + 4M(1-\delta)). \text{ (substituting } z'(1-q) \text{ for } z) \tag{8}$$

Notice $R(1,\delta) = 0$. Iterating inequality (8) for $z' = 1$ and then again for $z' = 1 - \underline{q}$ implies that $R(1-\underline{q},\delta) \leq \frac{4\rho}{\underline{z}l}4M(1-\delta)$ and $R((1-\underline{q})^2,\delta) \leq \frac{4\rho}{\underline{z}l}4M(1-\delta) + (\frac{4\rho}{\underline{z}l})^2 4M(1-\delta)$. More generally, for any $z \geq \underline{z}$

$$R(z,\delta) \leq 4M(1-\delta)\sum_{j=1}^{\bar{n}}\left(\frac{4\rho}{\underline{z}l}\right)^j$$

where $\bar{n}$ is the smallest integer such that $(1-\underline{q})^{\bar{n}} \leq \underline{z}$. Consequently, $\lim_{\delta\to 1} R(z,\delta) \leq \lim_{\delta\to 1} 4M(1-\delta)\sum_{j=1}^{\bar{n}}\left(\frac{4\rho}{\underline{z}l}\right)^j = 0$. □

4. Discussion

4.1. **The Stackelberg type.** In the repeated games that we consider, the dynamic Stackelberg strategy is not necessarily unique. For example in the game depicted in Figure 2, the grim trigger strategy is also a dynamic Stackelberg strategy. Mimicking the grim-trigger strategy would not however give player 1 a high payoff. This is because the punishment phase is also very costly for player 1. In contrast, the particular Stackelberg type that we choose is not very costly to mimic since the punishment phase is short, i.e., n^p is chosen minimally. If we had chosen any other finite length $n > n^p$ for the punishment phase, instead of n^p, our reputation result would still hold.

4.2. **Other commitment types.** A non-myopic player 2 may resist the Stackelberg type because she fears punishment for best responding or expects a reward for not best responding, either from another commitment type or from player 1's normal type. Our reputation result holds because, as we show, punishments or rewards cannot come from player 1's normal type; and we assume that the probability of another commitment type is small compared to the probability of the Stackelberg type. This restriction on the relative likelihood of other commitment types can be relaxed if the other commitment types are uniformly learnable. A uniformly learnable type reveals itself not to be the Stackelberg type, at a rate that is bounded away from zero, uniformly across all histories. If the other commitment types are uniformly learnable, then player 1 can play according to $\sigma_1(S)$ and ensure that player 2's posterior belief that player 1 is a type in Ω_- is arbitrarily small in finitely many periods. If player 2's posterior belief that player 1 is a type in Ω_- is small, then Theorem 1 implies that player 1's payoff is close to one, for sufficiently large discount factors. However, the restriction to uniformly learnable types is a non-trivial assumption. For example, it rules out the "perverse" type (see Schmidt (1993)) who plays like the dynamic Stackelberg type on the equilibrium path, but responds to deviations in a history dependent way.

In previous work, Schmidt (1993) and Celantani et al. (1996) establish reputation results with a non-mypopic player 2, even when the set of commitment types is arbitrary. Celantani et al. (1996) assume that player 2's moves are imperfectly observed with full support.[14] This assumption ensures that all relevant histories are sampled with positive probability, without any experimentation by player 2. If player 2's moves are imperfectly observed, then a rich set of commitment types are uniformly learnable. A similar assumption would also enable us to allow for a rich set of commitment types in the framework that we consider here.[15]

The reputation result of Schmidt (1993) obtains if the stage game is a game of conflicting interest, player 2's discount factor is fixed, and player 1 is arbitrarily more patient. Conflicting interests imply that the punishment that player 2 can expect from any other commitment type (her minimax payoff) is no worse than best responding to the Stackelberg type and receiving her minimax payoff. A commitment type may also reward player 2 for not best responding to the Stackelberg type. But, since player 2's discount factor is fixed, a reward for player 2 must entail behavior, that differs from the Stackelberg type, that occurs in a bounded number of periods T. If player 1 is sufficiently patient, he will mimic the Stackelberg type for these T periods, depriving player 2 from a reward and thus building a reputation. However, rewards for an equally patient player 2 need not accrue in a bounded number of periods. A commitment type that rewards player 2 for resisting the Stackelberg type, in a history dependent manner, can hinder player 1 from building a reputation against an equally patient opponent, even with a strictly conflicting interest stage game.

[14]Also, see Aoyagi (1996) for a similar assumption.

[15]See Atakan and Ekmekci (2008) which assumes player 2's moves are imperfect observed with full support and shows under this assumption that the set of other types can be taken as the set of all finite automata and the perfect information assumption can be dropped.

4.3. **Necessity of our assumptions.** In our analysis we assume perfect information and we restrict the set of stage games. Perfect information is necessary for a reputation result in locally non-conflicting interest stage games. Without perfect information, a folk theorem applies to the common interest game in Figure 3 (Cripps and Thomas (1997)), which is a locally non-conflicting interest game. For strictly conflicting interest stage games, perfect information is not required (see Cripps et al. (2005) or section 4.4).

Assumption 1 can fail in two ways. First, Assumption 1 fails if Γ is non-generic, i.e., the payoff profile where player 1 receives $\bar{g}_1$ is not unique in G. Such a failure is depicted in Figure 5. Second, Assumption 1 fails if, $(\bar{g}_1, \hat{g}_2) \in G$, but Γ is not a strictly conflicting interest game. Such a failure is depicted in Figure 6. A reputation result also fails to obtain in both of these examples.

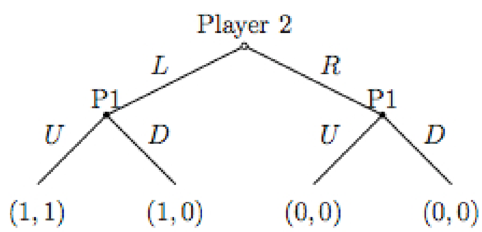

FIGURE 5. A non-generic common interest game.

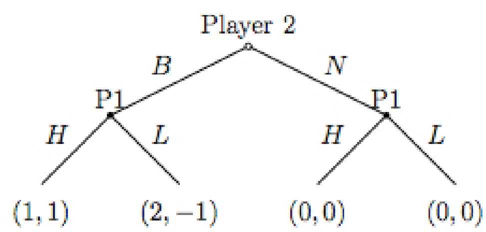

FIGURE 6. A moral hazard game that fails Assumption 1.

In the non-generic common interest game depicted in Figure 5 suppose that the Stackelberg type of player 1 always plays U and $\mu(S) < 1/2$. We describe a PBE where player 1 receives a payoff strictly lower than one. Suppose on the equilibrium path (R, U) is played in the first K periods and (L, U) is played thereafter. Player 1 does not build a reputation in this PBE. Choose K such that both players receive payoff equal to 1/2. Suppose, if player 2 deviates from equilibrium by playing L, then player 1's normal type reveals rationality by playing D, and the stage-game equilibrium (L, D) is played thereafter. Consequently, player 2 receives $\mu(S)$ if she deviates from the equilibrium strategy which is less than her equilibrium payoff 1/2.

In the moral hazard game depicted in Figure 6 player 1's dynamic Stackelberg payoff is 1.5 and player 2's minimax value is zero. In this game a dynamic Stackelberg strategy does not exist but there are strategies that deliver a payoff arbitrarily close to the dynamic Stackelberg payoff. Suppose that player 1's mixed actions are observed at the end of each period. One might conjecture that a payoff arbitrarily close to the dynamic Stackelberg payoff could be obtained by mimicking a Stackelberg type, S, that plays H with probability $1/2+\epsilon$. This is not the case. For example suppose on the equilibrium path player 1 plays H with probability $1/2+\epsilon$, in each period. Player 2 plays N for the first K periods and plays B thereafter. Choose K such that $\delta^K = 1/2$. Consequently, no reputation is built on the equilibrium path and equilibrium payoffs are $((1.5-\epsilon)/2, \epsilon/2)$. If player 1 deviates from equilibrium and reveals rationality, then player 2 plays N forever. If player 2 deviates from equilibrium and plays B, then player 1 reveals rationality by playing L. In the subsequent complete information game an equilibrium with payoffs $(1.5, 0)$ is played.[16] This construction is a PBE for any choice of ϵ, if $\mu(S) < 1/2$: If player 2 deviates and plays B, then she is facing S with probability $\mu(S)$ and receives payoff equal to ϵ, and she is facing the normal type with probability $1-\mu(S)$ and receives payoff equal to zero. However, $\mu(S)\epsilon < \epsilon/2$.

4.4. **Stage games of strictly conflicting interest.** Cripps, Dekel, and Pesendorfer (2005) obtain a reputation result for Bayes-Nash equilibria and simultaneous-move strictly conflicting interest stage games. A similar result can be obtained using the method developed here: Redefine $R(z,\delta)$ using Bayes-Nash equilibrium instead of PBE. The upper-bound established in Lemma 2 remains valid for Bayes-Nash equilibria. This is because all the arguments were constructed on an equilibrium path without any appeal to perfect information or subgame perfection. Also, $U_2(\sigma) \geq \hat{g}_2 = 0$ in any Bayes-Nash equilibrium. Consequently, the functional inequality (7) holds, and a reputation result follows.

Appendix A. Proof of Theorem 1

A.1. **Description of the argument.** The discussion in the main text was limited to $n^p = 1$, $\mu(\Omega_-) = 0$ and $R(z,\delta)$ non-increasing in z. We sketch the arguments used to relax these assumptions.

We extend the result to $n^p \geq 0$ in Lemma 5 and 6. Lemma 5 shows that player 2 faces an average per-period cost, $l > 0$, of not best responding to the dynamic Stackelberg type, i.e., $U_1(\sigma_1(S), \sigma_2, \delta) = 1 - r$ implies $U_2(\sigma_1(S), \sigma_2, \delta) \leq -lr$, if she is sufficiently patient. At any node where player 1 deviates from $\sigma_1(S)$, player 1 may have to carry-out an $n_p - 1$ period punishment

[16] Playing (N, L) in each period is a PBE of the complete information repeated game. Consequently, the threat of switching to (N, L) can incentivize a patient player 1 to play H with probability $1/2$ in each period.

phase, if he instead plays according to $\sigma_1(S)$. Lemma 6 generalizes Lemma 1 to account for this punishment phase.

Allowing for other commitment types with small probability requires incorporating the relative likelihood of the other commitment types as an additional state variable and accounting for the event that player 2 can be facing another commitment type, in the lower and upper-bound calculations. The relative likelihood $\frac{\mu(\Omega_-)}{\mu(S)}$ is non-increasing if player 1 plays the according to $\sigma_1(S)$. This monotonicity allows us to treat $\frac{\mu(\Omega_-)}{\mu(S)}$ as an additional state variable in Definitions 3 and 4. The effect of the other commitment types is at most $\pm M\phi$ on the lower-bound and the upper-bound. This is because player 1 is a commitment type with probability ϕ and player 2 can at most gain or loose M against any type. Consequently, if ϕ is small, then the effect of other commitment types on the functional equation is also small.

The central technical issue in the complete argument involves relaxing the assumption that $R(z,\delta)$ is non-increasing in z. Call z^* a *right-hand maximum* of R if $R(z,\delta) \leq R(z^*,\delta)$ for all $z > z^*$. If $z \in [\underline{z}, 1]$ and $z' > z$ are right-hand maximums of R, then the argument provided in the main text implies

$$R(z,\delta)(\underline{z}l - qC_1) \leq C_2 R(z',\delta) + C_3 M(1-\delta),$$

where $q = 1 - z/z'$; and C_1, C_2 and C_3 are positive constants independent of δ, z' and z that only depend on the parameters of the stage game as in (7). Rewriting,

$$q \geq \underline{z}D_1 - D_2\frac{R(z',\delta)}{R(z,\delta)} - D_3\frac{1-\delta}{R(z,\delta)},$$

where D_1, D_2 and D_3 are positive constants independent of δ, z' and z that only depend on the parameters of the stage game. We build a sequence of reputation levels that are "approximate" right-hand maximums of R. Let $K > 1$ be a constant such that $\underline{z}D_1 - D_2/K - D_3/K \geq \underline{z}D_1/2$. Let $z_n(\delta)$ be the supremum over reputation levels z such that $R(z,\delta) \geq K^n(1-\delta)$ (Definition 3). If z is greater than $z_n(\delta)$, then $R(z,\delta) < K^n(1-\delta)$. Each element of this sequence is "approximately" a right-hand maximum of R and we prove, for any $z_n(\delta) \in [\underline{z}, 1]$,

$$q_n(\delta) \geq \underline{z}D_1 - D_2\frac{R(z_{n-1}(\delta),\delta)}{R(z_n(\delta),\delta)} - D_3\frac{1-\delta}{R(z_n(\delta),\delta)} \geq \underline{z}D_1 - D_2\frac{K^{n-1}(1-\delta)}{K^n(1-\delta)} - D_3\frac{1-\delta}{K^n(1-\delta)},$$

where $q_n(\delta) = 1 - z_{n-1}(\delta)/z_n(\delta)$. Substituting in for K gives $q_n(\delta) \geq \underline{z}D_1/2 \equiv \underline{q}$. Let $\bar{n}$ denote the smallest integer such that $(1-\underline{q})^{\bar{n}} \leq \underline{z}$. So, $z_{\bar{n}}(\delta) \leq \underline{z}$ for any δ, and for any $z \geq \underline{z} \geq z_{\bar{n}}(\delta)$, $R(z,\delta) \leq K^{\bar{n}}(1-\delta)$. Consequently, $\lim_{\delta \to 1} R(z,\delta) = 0$.

A.2. **The proofs.**

Lemma 5. *Posit perfect information and Assumption 1. There exits $\delta^* \in [0,1)$ and $l > 0$ such that if $U_1(\sigma_1(S), \sigma_2, \delta) = 1 - r$, then $U_2(\sigma_1(S), \sigma_2, \delta) \leq -lr$, for all $\delta > \delta^*$.*

Proof. The definition of n^p given in inequality (4) implies that there exists a $\delta^* < 1$ and $l > 0$ such that, for all $\delta > \delta^*$,

$$g_2(a_1^s, a_2) + \sum_{k=1}^{n^p-1} \delta^k g_2(a_1^p, a_2') < -ln^p \tag{9}$$

for any $a_2 \in A_2$ such that $g_1(a_1^s, a_2) < 1$ and $a_2' \in A_2$. For public history $h^t = \{y^0, y^1, ..., y^t\}$, let $i(h^t) = 1$, if $g_1(y^t) < 1$ and $\sigma_1(S, h^t) = a_1^s$; and $i(h^t) = 0$, otherwise. Player 1 receives at least zero in any period t where $i(h^t) = 1$ and also receives at least zero in the subsequent $n^p - 1$ period punishment phase. In all other periods player 1 receives one. Consequently, $U_1(\sigma_1(S), \sigma_2, \delta) \geq 1 - n^p(1-\delta)\mathbb{E}_{(\sigma_1(S),\sigma_2)}\left[\sum_{t=0}^{\infty} \delta^t i(h^t)\right]$ and $(1-\delta)\mathbb{E}_{(\sigma_1(S),\sigma_2)}\left[\sum_{t=0}^{\infty} \delta^t i(h^t)\right] \geq r/n^p$.[17] If $i(h^t) = 1$, then player 2 receives a total discounted payoff of at most $-n^p l(1-\delta)$ for periods t through $t+n^p-1$, if $\delta > \delta^*$ by equation (9). In any period where a_1^s is played and $i(h^t)$=0 player 2 receives zero. Consequently, $U_2(\sigma_1(S), \sigma_2) \leq -n^p l(1-\delta)\mathbb{E}_{(\sigma_1(S),\sigma_2)}\left[\sum_{t=0}^{\infty} \delta^t i(h^t)\right] \leq -lr$, if $\delta > \delta^*$. □

In what follows, we assume that $\delta > \delta^*$; we fix $\underline{z} > 0$; $\gamma > 0$; $K > 1$ large such that

$$\frac{\underline{z}l}{2\rho} - \frac{1}{K} - \frac{2M(n^p+1))}{K^n} - \frac{2M}{\rho K^n} \geq \frac{\underline{z}l}{4\rho} > 0; \tag{10}$$

and we set $\max\{\phi, 1-\delta\} = \epsilon$.

Lemma 6. *Pick any PBE σ of $\Gamma^\infty(\mu,\delta)$, period t public history $h^t = (h^{t-1}, d_0)$, and suppose player 1 is to deviate from $\sigma_1(S)$ at node d_0 with positive probability given h^t. Let $h = (h^{t-1}, d)$ be the public history that is reached immediately (with positive probability under $\Pr_{(\sigma_1(N)|_{h^t}, \sigma_2|_{h^t})}$) following the deviation; let $h' = (h^{t-1}, d')$ be the public history that is reached immediately (with positive*

[17]The bound on player 1's payoff is crude especially for low δ.

probability under $\Pr_{(\sigma_1(S)|_{h^t},\sigma_2|_{h^t})}$*) if* $\sigma_1(S)$ *is used at* d *instead of deviating; and let* $\mu' = \mu(\cdot|h')$*, then* $U_1(\sigma,\delta|h) \geq 1 - R(\mu',\delta) - (n^p+1)M\epsilon$*. Consequently,*

$$|U_2(\sigma_1(N),\sigma_2,\delta|h)| \leq \rho(R(\mu',\delta) + (n^p+1)M\epsilon), \text{ if } \Gamma \text{ satisfies Ass. 1 (i), and}$$

$$U_2(\sigma_1(N),\sigma_2,\delta|h) \leq \rho(R(\mu',\delta) + (n^p+1)M\epsilon), \text{ if } \Gamma \text{ satisfies Ass. 1 (ii).}$$

Proof. If player 1 plays according to $\sigma_1(S)$ at d_0, then an $n^p - 1$ period punishment phase may follow. His payoff is at least zero in these periods. So, his payoff is at least $\delta^{n^p}(1 - R(\mu',\delta)) \geq 1 - R(\mu',\delta) - n^p M(1-\delta)$. Alternatively, he can deviate from $\sigma_1(S)$, receive $M(1-\delta)$ for the period, and $U_1(\sigma,\delta|h)$ is his continuation payoff. So, $M(1-\delta) + U_1(\sigma,\delta|h) \geq 1 - R(\mu',\delta) - n^p M(1-\delta)$ implies that $U_1(\sigma,\delta|h) \geq 1 - R(\mu',\delta) - (n^p+1)M\epsilon$. The bounds on player 2's payoff follow from equations (1) (2) and $(U_1(\sigma,\delta|h), U_2(\sigma_1(N),\sigma_2,\delta|h)) \in F$. □

Pick any period t public history $h^t = (h^{t-1}, d_0)$, and suppose player 1 moves at node d_0. Under perfect information, the public history that is reached immediately following h^t only depends on player 1's strategy and is independent from player 2's strategy. This distinction is relevant when we find a lower-bound for player 2's payoff in Lemma 8 by considering a non-equilibrium strategy for player 2 and applying Lemma 6.

Definition 3 (Reputation Thresholds)**.** *For each* $n \geq 0$*, let*

$$z_n(\delta,\phi) = \sup\{z : \exists \mu \in \Delta(\Omega) \text{ s.t. } R(\mu,\delta) \geq K^n\epsilon, \mu(S) = z, \frac{\mu(\Omega_-)}{\mu(S)} \leq \phi\}.$$

Definition 4. *For any* $\xi > 0$ *and* $z \in (0,1)$ *let*

$$\bar{R}(\xi,z,\delta,\phi) = \sup\{r : \exists \mu \in \Delta(\Omega) \text{ s.t. } R(\mu,\delta) \geq r, \mu(S) = z' \in [z-\xi, z], \frac{\mu(\Omega_-)}{\mu(S)} \leq \phi\}.$$

By definition, there exists μ such that $\mu(S) = z \in [z_n(\delta,\phi) - \xi, z_n(\delta,\phi)]$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$, and PBE σ of $\Gamma^\infty(\mu,\delta)$ such that σ_2 has resistance of at least $\bar{R}(\xi,z_n,\delta,\phi) - \xi$. Also, by definition, $\bar{R}(\xi,z_n,\delta,\phi) \geq K^n\epsilon$. The definition of $z_n(\delta,\phi)$ and $\bar{R}(\xi,z_n,\delta,\phi) \geq K^n\epsilon$ implies that if $\mu(S) \in [z_n(\delta,\phi) - \xi, z_{n-1}(\delta,\phi)]$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$, then $R(\mu,\delta) \leq \bar{R}(\xi,z_n,\delta,\phi)$ in any PBE profile σ of $\Gamma^\infty(\mu,\delta)$. The following lemma establishes an upper bound on Player 2's payoff in any equilibrium where the resistance is at least $\bar{R}(\xi,z_n,\delta,\phi) - \xi$.

Lemma 7. *Posit perfect information and Assumption 1. Pick* $\mu \in \Delta(\Omega)$ *such that* $\mu(S) = z \in [z_n(\delta,\phi)-\xi, z_n(\delta,\phi)]$ *and* $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$, *and pick PBE* σ *of* $\Gamma^\infty(\mu,\delta)$ *such that* $r(\delta,\sigma_2) \geq \bar{R}(\xi, z_n, \delta, \phi) - \xi$. *For the chosen PBE* σ,

$$U_2(\sigma,\delta) \leq \epsilon(\rho(\frac{q(\delta,\phi,n,\xi)\bar{R}(\xi,z_n,\delta,\phi)}{\epsilon} + K^{n-1} + 2M(n^p+1)) + 3M - \frac{(\bar{R}(\xi,z_n,\delta,\phi)-\xi)(z_n(\delta,\phi)-\xi)l}{\epsilon}), \tag{11}$$

where $q(\delta,\phi,n,\xi) = 1 - (z_n(\delta,\phi)-\xi)/z_{n-1}(\delta,\phi)$.

Proof. Choose pure strategy σ_2^* in the support of σ_2 such that $r(\sigma_2^*,\delta) \geq \bar{R}(\xi,z_n,\delta,\phi) - \xi$. Let profile $\sigma^* = (\{\sigma_1(\omega)\}_{\omega\in\Omega}, \sigma_2^*)$ and let $T = T(\sigma^*, \mu, q(\delta,\phi,n,\xi))$. Given that $\mu(S) = z$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$ if player 1 has not deviated from $\sigma_1(S)$ in h^t that is consistent with σ_2^*, then $\frac{\mu(\Omega_-|h^t)}{\mu(S|h^t)} \leq \phi$; and for $t \leq T$, $\mu(S|h^t) \geq z$; and for $t > T$, $\mu(S|h^t) \geq z_{n-1}$. We bound player 2's payoff in the events $\omega = N$ and $E_{[0,T)}$; $\omega = N$ and $E_{[T,\infty)}$; $\omega = S$; and $\omega \in \Omega_-$. Player 2's payoff until the time t that player 1 deviates from $\sigma_1(S)$ is at most $(1-\delta)M \leq \epsilon M$. Her payoff is zero if she plays a_2^b against a_1^s, Lemma 5 implies that her payoff is negative if she does not play a_2^b against a_1^s and a punishment phase is completed, and there can be at most one incomplete punishment phase until player 1 deviates from $\sigma_1(S)$.

Suppose that $h \in E_{[0,T)}$ and let $h^j = (h^{j-1}, d)$ denote the node where player 1 deviates from $\sigma_1(S)$ for the first time in the infinite public history h. Lemma 6 implies that $U_2(\sigma_1(N),\sigma_2,\delta|h^j) \leq \rho(\bar{R}(\xi,z_n,\delta,\phi) + \epsilon M(n^p+1))$. So,

$$U_2(\sigma_1(N),\sigma_2,\delta|E_{[0,T)}) \leq \epsilon M + \rho(\bar{R}(\xi,z_n,\delta,\phi) + \epsilon M(n^p+1)). \tag{12}$$

Suppose that $h \in E_{[T,\infty)}$ and let $h^j = (h^{j-1}, d)$ denote the node where player 1 deviates from $\sigma_1(S)$ for the first time in the infinite public history h. Lemma 6 implies that $U_2(\sigma_1(S),\sigma_2,\delta|h^j) \leq \rho\epsilon(K^{n-1} + M(n^p+1))$. So,

$$U_2(\sigma_1(N),\sigma_2,\delta|E_{[T,\infty)}) \leq \epsilon M + \rho\epsilon(K^{n-1} + M(n^p+1)). \tag{13}$$

Player 2 can get at most M against any other commitment type and this happens with probability $\phi z \leq \phi \leq \epsilon$. Player 2's resistance is $\bar{R}(\xi,z_n,\delta,\phi) - \xi$ in the equilibrium under consideration, she loses $(\bar{R}(\xi, z_n, \delta, \phi) - \xi)l$ against S by Lemma 5, and this happens with probability $z \geq z_n(\delta, \phi) - \xi$. The probability of N and $E_{[0,T)}$ is at most $q(\delta, \phi, n, \xi)$; and the probability of N and $E_{[T,\infty)}$ is at most one. Consequently, equations (12) and (13) imply

$$U_2(\sigma, \delta) \leq q(\delta, \phi, n, \xi)\rho\bar{R}(\xi, z_n, \delta, \phi) + \rho K^{n-1}\epsilon - (z_n(\delta, \phi) - \xi)(\bar{R}(\xi, z_n, \delta, \phi) - \xi)l + 2\rho\epsilon M(n^p + 1) + 3M\epsilon.$$

If $T = \infty$, then the bound is also valid. □

Lemma 8. *Posit perfect information and Assumption 1 item (i). Suppose that $\mu(S) = z \in [z_n(\delta, \phi) - \xi, z_n(\delta, \phi)]$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$. In any PBE σ of $\Gamma^\infty(\mu, \delta)$,*

$$U_2(\sigma, \delta) \geq -\epsilon \left(\rho \left(\frac{\bar{R}(\xi, z_n, \delta, \phi) q(\delta, \phi, n, \xi)}{\epsilon} + K^{n-1} + 3M(n^p + 1) \right) + M \right), \tag{14}$$

where $q(\delta, \phi, n, \xi) = 1 - (z_n(\delta, \phi) - \xi)/z_{n-1}(\delta, \phi)$.

Proof. Fix a PBE profile σ of $\Gamma^\infty(\mu, \delta)$ where $\mu(S) = z \in [z_n(\delta, \phi) - \xi, z_n(\delta, \phi)]$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$. Let σ_2^* denote a strategy that moves according to a_2^b after any period k public history h^k, if there is no deviation from $\sigma_1(S)$ in h^k, and coincides with PBE strategy σ_2 if player 1 has deviated from $\sigma_1(S)$ in h^k. Let profile $\sigma^* = (\{\sigma_1(\omega)\}_{\omega \in \Omega}, \sigma_2^*)$, let $T = T(\sigma^*, \mu, q(\delta, \phi, n, \xi))$. Player 2 receives zero in each period until player 1 deviates from $\sigma_1(S)$ because (a_1^s, a_2^b) is played under $(\sigma_1(S), \sigma_2^*)$. We apply Lemma 6 to obtain $U_2(\sigma_1(N), \sigma_2, \delta | E_{[0,T)}) \geq -\rho(\bar{R}(\xi, z_n, \delta, \phi) + \epsilon M(n^p + 1))$ and $U_2(\sigma_1(N), \sigma_2, \delta | E_{[T,\infty)}) \geq -\rho\epsilon(K^{n-1} + M(n^p + 1))$. Player 2 can get at least $-M$ against any other commitment type with probability at most $\phi \leq \epsilon$, gets zero against the Stackelberg type with probability z. Following the reasoning in Lemma 7 implies that

$$U_2(\sigma, \delta) \geq U_2(\sigma^*, \delta) \geq -\rho\bar{R}(\xi, z_n, \delta, \phi) q(\delta, \phi, n, \xi) - \rho\epsilon K^{n-1} - 2M\rho\epsilon(n^p + 1) - \epsilon M.$$

If $T = \infty$, then the bound is also valid. □

Proof of Theorem 1. If Γ satisfies Assumption 1 item (ii), then $U_2(\sigma, \delta) \geq \hat{g}_2 = 0$, and equation (14) is trivially satisfied. Combining the upper and lower bounds for $U_2(\sigma, \delta)$, given by equations (11) and (14), and simplifying by canceling ϵ delivers

$$(z_n(\delta, \phi) - \xi)l \frac{\bar{R}(\xi, z_n, \delta, \phi) - \xi}{\epsilon} \leq 2\rho \left(\frac{q(\delta, \phi, \xi)\bar{R}(\xi, z_n, \delta, \phi)}{\epsilon} + K^{n-1} + 2M(n^p + 1) \right) + 4M.$$

Let $q_n(\delta,\phi) = 1 - z_{n-1}(\delta,\phi)/z_n(\delta,\phi)$. $\bar{R}(\xi, z_n, \delta, \phi) \in [0,1]$ for each ξ, we pick any convergent subsequence and let $\lim_{\xi\to 0} \bar{R}(\xi, z_n, \delta, \phi) = \bar{R}(z_n, \delta, \phi)$. Taking $\xi \to 0$ implies that $q(\delta, \phi, n, \xi) \to q_n(\delta, \phi)$ and

$$z_n(\delta,\phi) l \bar{R}(z_n,\delta,\phi)/\epsilon \leq 2\rho(q_n(\delta,\phi)\bar{R}(z_n,\delta,\phi)/\epsilon + K^{n-1} + 2M(n^p+1)) + 4M.$$

Rearranging,

$$q_n(\delta,\phi) \geq \frac{z_n(\delta,\phi)l}{2\rho} - \frac{K^{n-1}\epsilon}{\bar{R}(z_n,\delta,\phi)} - \frac{2M(n^p+1)\epsilon}{\bar{R}(z_n,\delta,\phi)} - \frac{2M\epsilon}{\rho\bar{R}(z_n,\delta,\phi)}.$$

Also, $\bar{R}(\xi, z_n, \delta, \phi) \geq K^n\epsilon$ for each ξ implies that $\bar{R}(z_n, \delta, \phi) \geq K^n\epsilon$. Consequently,

$$q_n(\delta,\phi) \geq \frac{z_n(\delta,\phi)l}{2\rho} - \frac{K^{n-1}}{K^n} - \frac{2M(n^p+1)}{K^n} - \frac{2M}{\rho K^n}.$$

So, $q_n(\delta,\phi) \geq \frac{\underline{z}l}{2\rho} - \frac{1}{K} - \frac{2M(n^p+1)}{K^n} - \frac{2M}{\rho K^n}$, for any $z_n(\delta,\phi) \geq \underline{z}$. Substituting in our initial choice of K, defined in equation (10), implies $q_n(\delta,\phi) \geq \frac{\underline{z}l}{4\rho} \equiv \underline{q} > 0$, for any $z_n(\delta,\phi) \geq \underline{z}$. So, $z_n(\delta,\phi) \geq \underline{z}$ implies that $1 - z_n(\delta,\phi)/z_{n-1}(\delta,\phi) \geq \underline{q} > 0$ for all $\delta < 1$, $\phi > 0$ and $n = 0, 1, ..., \infty$. Then, for each $\delta < 1$ and $\phi > 0$, we have $z_{\bar{n}(\underline{q})}(\delta,\phi) < \underline{z}$, where $\bar{n}(\underline{q})$ is the smallest integer j such that $(1-\underline{q})^j < \underline{z}$. By definition, if $\mu(S) \geq z_{\bar{n}(\underline{q})}(\delta,\phi)$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$, then $R(\mu,\delta) \leq K^{\bar{n}(\underline{q})}\epsilon$. Consequently, if $\max\{1-\delta,\phi\} = \epsilon < \frac{\gamma}{K^{\bar{n}(\underline{q})}}$, $\mu(S) \geq \underline{z}$ and $\frac{\mu(\Omega_-)}{\mu(S)} \leq \phi$, then $U_1(\sigma,\delta) > 1-\gamma$ for all PBE σ of $\Gamma^\infty(\mu,\delta)$. □

References

Aoyagi, M. (1996): "Reputation and dynamic Stackelberg leadership in infinitely repeated games," *Journal of Economic Theory*, 71(2), 378–393.

Atakan, A., and M. Ekmekci (2008): "Reputation with long run players and imperfect observation," *Working paper.*

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.

EVANS, R., AND J. THOMAS (1997): "Reputation and Experimentation in Repeated Games With Two Long-Run Players," *Econometrica*, 65(5), 1153–1173.

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.

MAILATH, G., AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, Oxford, UK.

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

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