Atakan, Alp E.; Ekmekci, Mehmet

Working Paper

Reputation with long run players and imperfect

Discussion Paper, No. 1507

Provided in Cooperation with:
Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Atakan, Alp E.; Ekmekci, Mehmet (2009) : Reputation with long run players and imperfect, Discussion Paper, No. 1507, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:
https://hdl.handle.net/10419/59645

REPUTATION IN LONG-RUN RELATIONSHIPS

ALP E. ATAKAN AND MEHMET EKMEKCI

Abstract. We model a long-run relationship as an infinitely repeated game played by two equally patient agents. In each period, the agents play an extensive-form game of perfect information. There is incomplete information about the type of player 1 while player 2's type is commonly known. We show that a sufficiently patient player 1 can leverage player 2's uncertainty about his type to secure his highest payoff in any perfect Bayesian equilibrium of the repeated game.

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

JEL Classification Numbers: C73, D83.

Date: December, 2009.

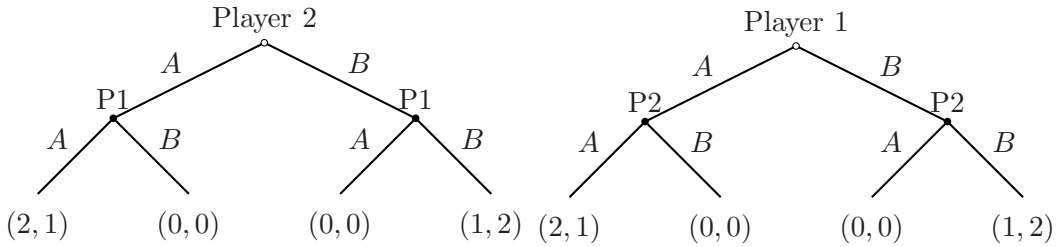
We thank Nabil Al-Najjar, Martin Cripps, Eddie Dekel, Drew Fudenberg, Christoph Kuzmics, George Mailath and Larry Samuelson for their comments; and Umberto Garfagnini for research assistance. This paper was previously circulated under the title "Reputation with Long-run Players."

Contents

1. Introduction and Related Literature

Maintaining a reputation can benefit an economic agent since it lends credibility to his future commitments, threats or promises. So, the agent may forego short-run profits to cultivate a reputation in anticipation of long-run benefits. The incentive to build a reputation is most pronounced if the agent is patient, that is, if the short-run loss is less important to the agent than the long-run benefit. There is a tension, however, if the agent faces an opponent who is equally patient: the opponent may also sacrifice short-run payoffs to extensively test the agent's resolve to go through with his commitments, threats or promises. This can make it prohibitively expensive and undesirable to build a reputation. In this paper we focus on equally patient agents to highlight this tension. We explore how reputation concerns affect the outcomes of repeated interactions between two *equally patient* agents who make choices sequentially.

Consider two players involved in a long-run relationship, for example a husband and wife, an employee and an employer, two countries, or two legislators. In each period of the relationship the two players must decide whether to undertake policy A, policy B or to undertake neither of the two policies. Unanimity is required for any policy to be chosen. Player 1 (he) prefers policy A, player 2 (she) prefers policy B, and both players prefer some policy to no policy at all. These policies can represent competing treaties in a pollution abatement negotiation between two countries, budget alternatives under consideration for two political rivals, or even weekend plans bargained over by a married couple. The games depicted in Figure 1a and Figure 1b are possible versions of the strategic situation that the players repeatedly face.[1]



(A) Player 2 moves first. (B) Player 1 moves first.

Figure 1. The battle of the sexes.

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

Suppose that player 2 suspects that her opponent always chooses policy A. In particular, she believes that player 1 is either fully rational, or with probability $z > 0$, a Stackelberg type who is committed to playing A. A rational player 1, cognizant of player 2's uncertainty, has an incentive to mimic the Stackelberg type. If player 2 is convinced that player 1 is the Stackelberg type, then she will have no choice but to play A and policy A will be the outcome in each period. So, a patient player 1 may play A for many periods, even if player 2 plays B, (i.e., at the expense of reaching an agreement) in order to convince player 2 that he is indeed the Stackelberg type. However, player 2 knows that player 1 has an incentive to mimic the Stackelberg type. Consequently, an equally patient player 2 may play B (i.e., resist playing A) for many periods making building a reputation particularly costly, especially if she deems it sufficient likely that player 1 is rational and he will eventually start playing B.

Given these two opposing forces, can player 1 build a reputation and ensure that policy A is implemented? Or alternatively, will screening by player 2 keep a rational player 1 from building a reputation? Our main finding addresses these questions: Suppose that the players are equally and arbitrarily patient. If the game in Figure 1a or Figure 1b is played repeatedly, then policy A is implemented in each period and player 1 receives a payoff equal to 2 in any perfect Bayesian equilibrium of the repeated game. This outcome is independent of which player moves first and independent of how small, z, the initial uncertainty about player 1's type is.

More generally, we model a long-run relationship as an infinitely repeated game played by two equally patient agents. In each period, the agents play an extensive-form game of perfect information. We assume that player 2 is uncertain about the type of player 1 while player 1 is perfectly informed about the type of player 2.

In the previous example, player 1's reputation allowed him to credibly commit to always choosing the same action. However we can conceive of other strategic situations where player 1 may want to commit to a more complex strategy that rewards or punishes his opponent in a history dependent way. For example, player 1 may want to be known for playing tit-for-tat, or for punishing bad behavior consistently. To capture reputation effects more generally, we assume that player 1 is either fully rational or one of many *commitment* types. Each commitment type is programmed to play a certain repeated game strategy. The commitment type central to our analysis is the *dynamic*

Stackelberg type. This type plays the repeated game strategy that player 1 would choose, if player 1 could publicly pre-commit to any repeated game strategy. Ideally, player 1 would like to convince his opponent that his future actions will fully conform to the behavior of the dynamic Stackelberg type.

In this framework, we prove a *reputation result* that uniquely characterizes the long-run outcome. We show that a sufficiently patient player 1 can use his ability to mimic the dynamic Stackelberg type and his opponent's uncertainty about his type to secure his most preferred outcome for the repeated game. More precisely, if player 1 is a dynamic Stackelberg type with positive probability, then player 1 receives his highest payoff that is consistent with the individual rationality of player 2, in any perfect equilibrium of the repeated game, as the common discount factor of the players converges to one.

1.1. **Related literature.** This paper is closely related to the literature on reputation effects in repeated games. We make three main contributions to this literature: First, we provide a reputation result for a new class of repeated games played by two equally patient players: repeated extensive-form games of perfect information. Previous reputation results for equally patient agents are for only a limited class of repeated simultaneous move games. Most previous work has instead focused on a long-run (i.e, patient) player facing a myopic (i.e, infinitely less patient) opponent. Second, we highlight the distinct role that perfect information plays for a reputation result with equally patient agents. Third, we introduce novel methods, inspired by the bargaining literature (Myerson (1991), section 8.8), to analyze reputation effects in repeated games.

Much of the previous literature on reputation considers a patient player 1 who faces a myopic opponent. Most prominently, Fudenberg and Levine (1989, 1992) showed that if there is positive probability that player 1 is a type committed to playing the Stackelberg action in every period, then player 1 gets at least his static Stackelberg payoff in any equilibrium of the repeated game.[2] 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 (see Schmidt (1993), Celantani et al. (1996), Aoyagi (1996), or Evans and Thomas (1997)). Again, however, the repeated interactions

[2]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.

that these papers consider are genuinely long-run only from the point of view of player 1 and this feature is crucial for the results.

In a game with a non-myopic opponent, player 1 may achieve a payoff that exceeds his static Stackelberg payoff with a history dependent strategy that rewards or punishes player 2. Conversely, future punishments or rewards can induce player 2 to not best respond to a Stackelberg action and thereby force player 1 below his static Stackelberg payoff.[3] These complications render reputation effects fragile in repeated games with equally patient players: A reputation result obtains in a repeated simultaneous-move game only if there is a strictly dominant action in the stage game (Chan (2000)), or if there are strictly conflicting interests in the stage game (Cripps et al. (2005)).[4] For other repeated simultaneous-move games any individually rational payoff can be sustained in a perfect equilibrium, if the players are sufficiently patient (see folk theorems by Cripps and Thomas (1997) and Chan (2000)).

Previous literature on reputation with equally patient agents focuses on repeated simultaneous-move games (e.g., Cripps and Thomas (1995), Cripps and Thomas (1997), Chan (2000) or Cripps et al. (2005)).[5] In contrast, we focus on repeated extensive-form games of perfect information. This allows us to establish a reputation result that covers a wide class of games. In particular, we establish reputation results for repeated *locally non-conflicting interests* games, and repeated strictly conflicting interests games.[6] 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 is played simultaneously, then a folk theorem applies to games with locally non-conflicting interests even under incomplete information (see Cripps and Thomas (1997) or Chan (2000)).

[3]Player 2 may expect punishments or rewards from either the rational type of player 1 after he chooses a move that would not be chosen by the Stackelberg type (Celantani et al. (1996) section 5 or Cripps and Thomas (1997)), or from a commitment type other than the Stackelberg type (Schmidt (1993) or Celantani et al. (1996)).

[4]There are strictly conflicting interests in a game if the action which is the best for player 1 is the worst for his opponent. See Assumption 1 for a precise statement.

[5] Cripps and Thomas (1997), Chan (2000) Cripps et al. (2005) assume equal discount factors, as we do, to model equally patient agents whereas Cripps and Thomas (1995) assumes no discounting and uses the limit of means. Also see Bar-Isaac (2003), Cripps and Thomas (2003), Wiseman (2005), Faingold (2005) and Peski (2008) for related results in repeated games with one-sided incomplete information

[6]There are locally non-conflicting interests in a game if the payoff profile where player 1 receives his highest payoff is strictly individually rational for player 2.

Our finding points out that reputation effects are particularly salient in repeated sequential-move games (i.e., games of perfect information), whereas reputation effects are absent in a wide range of repeated simultaneous-move games. For example, our reputation result implies a unique outcome for the sequential-move battle of the sexes (Figure 1a or Figure 1b) whereas in the simultaneous-move battle of the sexes game (Figure 2) a folk theorem obtains. For a more striking example

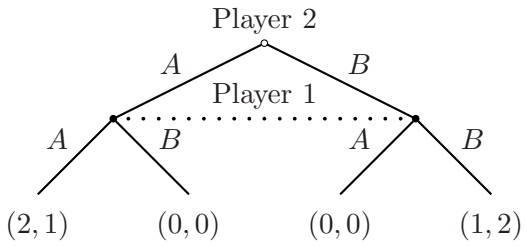
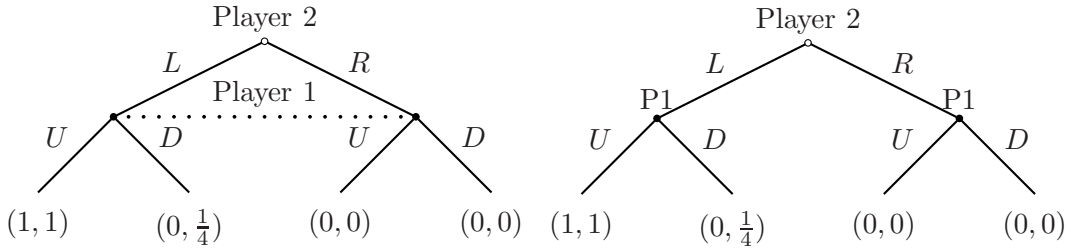


FIGURE 2. The simultaneous-move battle of the sexes.

consider the repeated common interest game (Figure 3a or Figure 3b), where player 1 is potentially a Stackelberg type who always plays U. This game appears as a strong candidate for reputation effects to arise. It is costless for player 1 to mimic the Stackelberg type and build a reputation. Also, player 2 unambiguously benefits if player 1 is able to build a reputation and concentrate play on (U, L). Surprisingly, any individually rational payoff profile can be sustained in a perfect Bayesian equilibrium, if the players are arbitrarily patient (Cripps and Thomas (1997)). In contrast, in the repeated sequential-move game, the players receive a payoff equal to one in any perfect Bayesian equilibrium. We discuss in detail the Cripps and Thomas (1997) construction for the simultaneous-move common interest game, and how perfect information (sequentiality) allows us to avoid their conclusion in section 4.1.



(A) Simultaneous-move. (B) Sequential-move.

FIGURE 3. A common interest game.

With two equally patient players the techniques of Fudenberg and Levine (1989, 1992), which are commonly used to establish reputation results, are not applicable. 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's normal type reveals rationality. Subgame perfection, coupled with perfect information, imposes tight bounds on player 2's continuation payoffs at these nodes. These bounds preclude the possibility that player 1 builds a reputation slowly and punishes player 2 for best responding to the Stackelberg strategy.

1.2. **Outline of the paper.** The paper proceeds as follows: section 2 describes the repeated game of complete and incomplete information. Section 3 presents the main reputation result and details the argument in an example. The proof of the main theorem is described in section 3.1 and can be found in the appendix. Section 4 discusses our assumptions, the results and an extension. Specifically, sections 4.1 and 4.2 show the necessity of our assumptions through some examples. Section 4.3 expands further on the dynamic Stackelberg type. Section 4.4 considers issues pertaining to commitment types other than the dynamic Stackelberg type and learning. Finally, section 4.5 argues that our reputation result obtains for strictly conflicting interest stage games even without the perfect information assumption.

2. The Model

In the repeated game 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)**.

D is the set of nodes of the stage game Γ (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 $\mathcal{A}_i$.[7] 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. Slightly abusing notation

[7]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.

we let $g_i(a) = g_i(y(a))$ for any $a \in A_1 \times A_2$, and we let $g_i(\alpha)$ denote the payoff to mixed action profile $\alpha \in \mathcal{A}_1 \times \mathcal{A}_2$.

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

Types and Strategies. Before time 0 nature selects player 1's 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.

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.[8] 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).

The repeated game and payoffs. A player's repeated game payoff is the normalized discounted sum of the stage game payoffs. For any infinite public history $h^\infty = \{y^0, y^1, ..., \}$, $u_i(h^\infty, \delta) = (1-\delta)\sum_{k=0}^{\infty} \delta^k g_i(y^k)$, 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, are given by $U_1(\sigma, \delta | h^t) = U_1(\sigma_1(N), \sigma_2, \delta | h^t)$ and

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

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, $U_i(\sigma, \delta) = U_i(\sigma, \delta | h^0)$.

[8] Abusing notation we will use σ_i to also denote mixed repeated game strategies for player i. Behavior strategies and mixed strategies are equivalent, by Kuhn's theorem, in this framework.

The repeated game of complete information, that is, the repeated game without any commitment types, with discount factor equal to $\delta \in [0,1)$, is denoted $\Gamma^\infty(\delta)$. The repeated game of incomplete information, with the prior over the set of commitment types given by $\mu \in \Delta(\Omega)$ and the discount factor equal to $\delta \in [0,1)$, is denoted $\Gamma^\infty(\mu,\delta)$.

The stage game. The minimax payoff for player i is $\hat{g}_i = \min_{\alpha_j \in \mathcal{A}_j} \max_{\alpha_i \in \mathcal{A}_i} g_i(\alpha_i, \alpha_j)$. For games that satisfy perfect information there exists $a_1^p \in A_1$ such that $g_2(a_1^p, a_2) \leq \hat{g}_2$ for all $a_2 \in A_2$.[9] The set of feasible payoffs is $F = co\{g_1(a_1,a_2), g_2(a_1,a_2) : (a_1,a_2) \in A_1 \times A_2\}$; and the set of feasible and individually rational payoffs is $G = F \cap \{(g_1,g_2) : g_1 \geq \hat{g}_1, g_2 \geq \hat{g}_2\}$. Let $\bar{g}_1 = \max\{g_1 : (g_1,g_2) \in G\}$. We assume that the stage game satisfies the following assumption.

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

(*i*) *(Locally Non-Conflicting Interests) 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 Interests) 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 $(\bar{g}_1, g_2) \in G$.*[10]

Assumption 1 requires that the payoff profile where player 1 obtains $\bar{g}_1$ is unique (for example, this is true if the game Γ is a *generic* extensive form game). 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. The sequential games that we discussed in the introduction (Figures 1a, 1b, and 3b) all satisfy Assumption 1. For example, the sequential battle of the sexes where player 1 moves second (Figure 1a) and where player 1 moves first (Figure 1b) are games of locally non-conflicting interests and strictly conflicting interests, respectively. We discuss some games that do not satisfy Assumption 1 in section 4.2.

If there is an action for player 1, $a_1 \in A_1$, and a best response for player 2 to action a_1, $a_2 \in A_2$, such that $g_1(a_1,a_2) = \bar{g}_1$, then we define $a_1^s = a_1$ and $a_2^b = a_2$. Otherwise, we define

[9] 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 $g_2(a_1^p, a_2) = \hat{g}_2$.

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

$(a_1^s, a_2^b) \in A_1 \times A_2$ as an action profile such that $g_1(a_1^s, a_2^b) = \bar{g}_1$.[11] For example, in the battle of the sexes game depicted in Figure 1a the stage game action a_1^s chooses move A in both of player 1's information sets and $a_2^b = A$. In this game a_1^s and a_2^b are uniquely determined. For the common interest game depicted in Figure 3b the action a_2^b is unique and $a_2^b = L$. However, in this game a_1^s is not uniquely determined. It can either be the stage game action that always chooses U or the action that chooses U after L and D after R. In this case we pick a_1^s arbitrarily from these two choices.[12]

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

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

The set of feasible payoffs in the repeated game is equal to the set of feasible stage game payoffs F. If Γ satisfies Assumption 1 (i), then equation (1) implies that

$$\left| \frac{U_2(\sigma_1, \sigma_2, \delta) - g_2(a_1^s, a_2^b)}{\bar{g}_1 - U_1(\sigma_1, \sigma_2, \delta)} \right| \leq \rho, \text{ for any two repeated game strategies } \sigma_1 \in \Sigma_1 \text{ and } \sigma_2 \in \Sigma_2.$$

If Γ satisfies Assumption 1 (ii), then equation (2) implies that

$$U_2(\sigma_1, \sigma_2, \delta) - g_2(a_1^s, a_2^b) \leq \rho(\bar{g}_1 - U_1(\sigma_1, \sigma_2, \delta)), \text{ for any two repeated game strategies } \sigma_1 \in \Sigma_1 \text{ and } \sigma_2 \in \Sigma_2.$$

Dynamic Stackelberg payoff, strategy and type. Let

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

[11] 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 there is more than one action profile that satisfies our definition, then we pick (a_1^s, a_2^b) arbitrarily as any one of these action profiles.

[12] If Γ is a strictly conflicting interest game, then a_2^b is a best response to a_1^s. If Γ is a locally non-conflicting interest game, then a_2^b is not necessarily a best response to a_1^s. For an example that satisfies Assumption 1 where a_2^b is not a best response to a_1^s see Figure 4.

where $BR(\sigma_1, \delta)$ denotes the set of best responses of player 2, in the repeated game $\Gamma^\infty(\delta)$, 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.[13] The dynamic Stackelberg payoff for player 1 is the highest payoff that player 1 can secure in the repeated game through public pre-commitment to a repeated game strategy. The dynamic Stackelberg strategy for player 1 is a repeated game strategy such that any best response of player 2 to this strategy gives player 1 at least his dynamic Stackelberg payoff.

If Γ satisfies Assumption 1, then $U_1^s(\delta) = \bar{g}_1$ 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)).

We focus on a particular dynamic Stackelberg type, denoted S, that plays a strategy $\sigma_1(S)$. The strategy $\sigma_1(S)$ has a profit 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 $g_1(y_{t-1}) = \bar{g}_1$. The strategy moves to the punishment phase in period t, if it was in the profit phase in period $t-1$ and $g_1(y_{t-1}) \neq \bar{g}_1$. 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 $\bar{g}_1$. The number of punishment periods $n^p - 1$ is the smallest integer such that

$$g_2(a_1^s, a_2) + (n^p - 1)\hat{g}_2 < n^p g_2(a_1^s, a_2^b) \tag{3}$$

for any $a_2 \in A_2$ such that $g_1(a_1^s, a_2) < g_1(a_1^s, a_2^b) = \bar{g}_1$. 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

[13]The terminology follows Aoyagi (1996) and Evans and Thomas (1997).

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 $U_1(\sigma_1(S), \sigma_2, \delta) = \bar{g}_1$, 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 1a, Figure 1b or Figure 3b). If $n^p > 1$, then the dynamic Stackelberg payoff strictly exceeds the static Stackelberg payoff for a sufficiently high discount factor (see Figure 4).

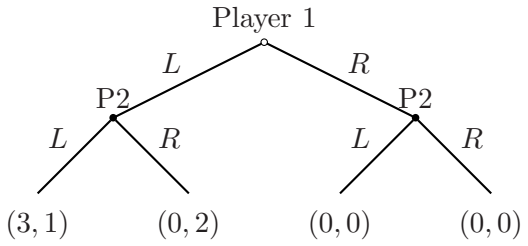


FIGURE 4. For this stage game $a_1^s = L$ and we pick a_2^b as the action that always chooses move L. Also, in this example the Stackelberg type S plays L in the profit phase, R in the two period punishment phase of the repeated game and $n^p = 3$.

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

Equilibrium and beliefs. The 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. 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

Our 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 $\mu(S) > 0$. The theorem provides a lower bound on player 1's payoff in any PBE and the formal statement is given below.

Theorem 1. *Assume perfect information and Assumption 1. For any $\delta \in [0,1)$, any $\mu \in \Delta(\Omega)$ such that $\mu(S) > 0$ and any PBE strategy profile σ of $\Gamma^{\infty}(\mu,\delta)$*

$$U_1(\sigma,\delta) \geq \bar{g}_1 - f(\underline{z}) \max\{1-\delta, \mu(\Omega_-)\},$$

where $\underline{z} = \mu(S)$ and f is a decreasing, positive function that is independent of δ and μ.

Proof. The function f is defined in equation (9) in the appendix. The proof is in the appendix. □

The theorem implies that as δ goes to one and $\mu(\Omega_-)$ (the probability of other commitment types) goes to zero, player 1's payoff converges to $\bar{g}_1$, his highest payoff. Consequently, a normal type for player 1 can secure a payoff arbitrarily close to $\bar{g}_1$, his dynamic Stackelberg payoff, in any PBE of the repeated game, for a sufficiently high discount factor and for sufficiently low probability mass on other commitment types. Player 1 can attain the bound given in the theorem by simply mimicking the Stackelberg type. Notice that the bound given in the theorem is not particularly sharp, if the probability of other commitment types, $\mu(\Omega_-)$, is substantial. However, under certain assumptions player 1 can receive a payoff arbitrarily close to $\bar{g}_1$, with no restrictions on the probability of other commitment types. We discuss such issues related to other commitment types in section 4.4.

In this section we prove the reputation result given in Theorem 1, under the assumption that $\mu(\Omega_-) = 0$, for the example depicted in Figure 5. At the end of the section we discuss the main argument for Theorem 1 that is given in the Appendix.

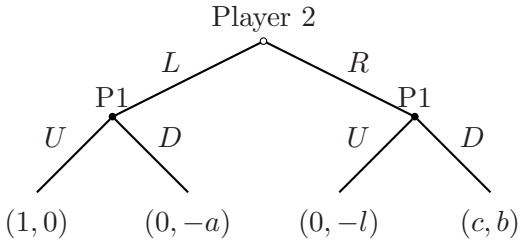


FIGURE 5. Assume that $l \in (0,1]$, $a \in (0,1]$, $b \in [-1,1]$ and $c \in [0,1/2]$. This is a game of locally non-conflicting interests where $n^p = 1$. If $l = 1$, $a = 1$, $b = 1$ and $c = 1/2$, then this is a normalized version of the sequential-move battle of the sexes game depicted in Figure 1a. If $l = 1$, $a = 3/4$, $b = -1$ and $c = 0$, then this is a normalized version of the sequential-move common interest game depicted in Figure 3b.

Recall $(a_1^s, a_2^b) \in A_1 \times A_2$, defined in the text preceding equation (1), is an action profile such that $g_1(a_1^s, a_2^b)$ is equal to player 1's highest stage game payoff. For this example, player 1's highest

stage game payoff is equal to one, the stage game action a_1^s plays U after either L or R; and a_2^b is a best response to a_1^s and plays L. Also, $n^p = 1$ and the Stackelberg type S plays a_1^s in each period of the repeated game (i.e, S plays U at each decision node of player 1).[14] Our reputation result, for this particular example, is stated below.

Corollary 1. *Suppose that the stage game* Γ *is given by Figure 5 and assume that* $\mu(\Omega_-) = 0$. *For any reputation level* $\mu(S) = \underline{z} > 0$ *we have* $\lim_{\delta \to 1} U_1(\sigma(\delta), \delta) = 1$ *where* $\sigma(\delta)$ *is a PBE strategy profile for the repeated game* $\Gamma^\infty(\mu, \delta)$.

In what follows, because $\mu(\Omega_-) = 0$, we use $z \in [0, 1]$ to represent the measure μ. One should understand this to mean $\mu(S) = z$ and $\mu(N) = 1 - z$. We begin with some definitions. Let the resistance of strategy σ_2 be given by

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

Notice the definition of the resistance of strategy σ_2 implies that the expected discounted number of periods where player 2 plays R against U is $r(\sigma_2, \delta)$, if player 2 uses strategy σ_2 and player 1 uses strategy $\sigma_1(S)$. Consequently, if player 2 uses strategy σ_2 and her opponent uses strategy $\sigma_1(S)$, then player 2's payoff, $U_2(\sigma_1(S), \sigma_2, \delta)$, is equal to $-lr(\sigma_2, \delta)$.

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

Definition 1 (Resistance function)**.** *For any measure* $z > 0$ *and* $\delta \in [0, 1)$ *let*

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

In this section we work under the hypothesis that $R(z, \delta)$ *is a non-increasing function of* z *for each* $\delta \in [0, 1)$. This is for expositional convenience only and allows us to convey the main ideas of the argument without the more technical details. The main proof, given in the appendix, does not use this assumption, and the steps involved in relaxing this assumption are discussed in section 3.1.

At the start of any period t, if player 1's reputation level is at least $z > 0$, then player 1 can guarantee a continuation payoff of at least $1 - R(z, \delta)$, by playing according to the Stackelberg

[14] For the definition of n^p see equation (3). For this game the static and dynamic Stackelberg payoffs coincide and are equal to one, for any discount factor.

strategy $\sigma_1(S)$. This follows from the definition of R, subgame perfection, and our assumption that R is non-increasing. We will argue that $\lim_{\delta\to 1} R(\underline{z},\delta)=0$, for any $\underline{z}>0$.

Consider a PBE σ of the repeated game $\Gamma^\infty(z,\delta)$. Suppose that the players are at a history in which player 1 has played U in each period before t and player 2 has played $a_2\in\{L,R\}$ in period t. Further suppose that player 1 plays D with positive probability at this decision node, i.e., player 1 reveals that he is not the Stackelberg type. Also, let player 1's reputation level be $z'>0$ at the start of period $t+1$, if he plays U instead of D. In the next lemma we bound the continuation payoffs for both players in terms of resistance $R(z',\delta)$ at any such decision node. The argument for the lemma is as follows: if player 1 is playing D with positive probability, then the payoff from playing D must be at least as large as the payoff from playing U. However, if player 1 plays U, he gets at worst zero for the period, ensures that his reputation is z' at the start of the subsequent period, and thus guarantees $1-R(z',\delta)$ at the start of period $t+1$. Given this lower bound on player 1's continuation payoff a bound on player 2's continuation payoff follows from equation (1).

Lemma 1. *Suppose $z>0$ and pick any PBE σ of $\Gamma^\infty(z,\delta)$, period t public history h^t where player 1 has played U in each period, and suppose player 1 is to play D in period t given history (h^t,a_2), where $a_2\in\{L,R\}$. Let $z'=\mu(S|h^t,a_2,U)$; then $|U_2(\sigma_1(N),\sigma_2,\delta|h^t,a_2,D)|\leq\rho(R(z',\delta)+(1-\delta)/\delta)$, where $\rho\leq 1$.*

Proof. If player 1 plays U in period t, then his reputation level is $z'=\mu(S|h^t,a_2,U)$ and he can guarantee a 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 U. Consequently, his payoff from playing U is at least $\delta U_1(\sigma,\delta|h^t,a_2,U)\geq\delta(1-R(z',\delta))$. If instead player 1 plays D, then he can get at most c for the current period and $\delta U_1(\sigma,\delta|h^t,a_2,D)$ as his continuation payoff. Because player 1 is willing to play D instead of U we have $(1-\delta)c+\delta U_1(\sigma,\delta|h^t,a_2,D)\geq\delta U_1(\sigma,\delta|h^t,a_2,U)$. Hence, $U_1(\sigma,\delta|h^t,d^t,D)\geq 1-R(z',\delta)-(1-\delta)c/\delta\geq 1-R(z',\delta)-(1-\delta)/\delta$. The bound on player 2's payoff follows from equation (1) and $(U_1(\sigma,\delta|h^t,d^t,D),U_2(\sigma_1(N),\sigma_2,\delta|h^t,d^t,D))\in F$. The constant ρ in equation (1) is equal to 1 for this particular game. The argument is also depicted graphically in Figure 6. □

We now use Lemma 1 to sketch the argument for Corollary 1. Suppose that player 1's reputation level is z. Consider a PBE $\sigma=(\sigma_1(N),\sigma_1(S),\sigma_2)$ where player 2 resists the Stackelberg type by

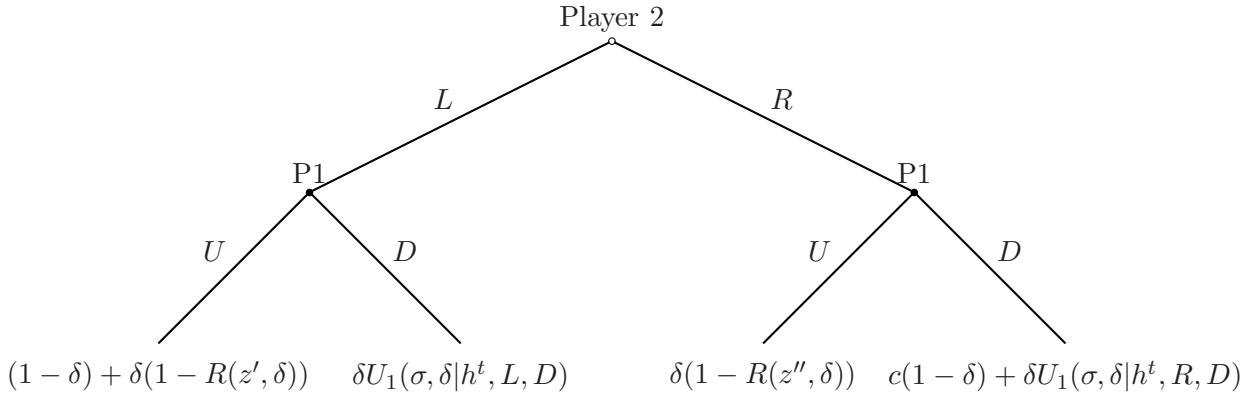


FIGURE 6. This figure depicts the payoff player 1 can guarantee by playing U and his payoff if he instead plays D. In the figure $z' = \mu(S|h^t, L, U)$ and $z'' = \mu(S|h^t, R, U)$. If player 1 is to play D after R, then $(1-\delta)c + \delta U_1(\sigma, \delta|h^t, R, D) \geq \delta(1 - R(z'', \delta))$. Consequently, $U_1(\sigma, \delta|h^t, R, D) \geq 1 - R(z'', \delta) - (1-\delta)c/\delta$, equation (1) and $\rho \leq 1$ implies that $|U_2(\sigma, \delta|h^t, R, D)| \leq R(z'', \delta) + (1-\delta)c/\delta$. Similarly, if player 1 is to play D after L, then $|U_2(\sigma, \delta|h^t, L, D)| \leq R(z', \delta) - (1-\delta)/\delta$.

approximately $R(z, \delta)$. In this PBE player 2 loses approximately $lR(z, \delta)$ in the event that player 1 is the Stackelberg type. We compare player 2's payoff in this PBE with her payoff if she uses an *alternative strategy* that plays L until player 1 plays D for the first time and then reverts back to the equilibrium strategy σ_2. If player 2 uses the alternative strategy, then she avoids losing $lR(z, \delta)$ in the event that player 1 is the Stackelberg type. We then use the fact that the PBE strategy σ_2 must give player 2 a payoff that is at least as great as the payoff from using the alternative strategy. This establishes a bound on $R(z, \delta)$, for any z sufficiently close to 1.

Suppose that player 1 plays D for the first time in some period t. In each period, up to period t player 2 receives at best zero, in period 2 she receives at best $1-\delta$; and she receives at most $R(z, \delta) + (1-\delta)/\delta$ as a continuation payoff after period t, by Lemma 1 and by our assumption that R is non-increasing. Consequently, player 2 gets at most $\delta^t(1-\delta)+\delta^{t+1}(R(z, \delta)+(1-\delta)/\delta) \leq R(z, \delta)+2(1-\delta)$, if player 1 plays D for the first time in period t. If player 1 always plays U in each period, then player 2 receives at most $-lR(z, \delta)$. Player 1 will play U in every period with at least probability z because type S always plays U. So, player 1 will play D in some period t, with probability at most $1 - z$. Thus, player 2's payoff in PBE σ is at most $(1-z)(R(z, \delta) + 2(1-\delta)) - zlR(z, \delta)$. Lemma 2, that we state below, establishes an upper bound that formalizes this line of reasoning.

Suppose that player 2 uses the alternative strategy and player 1 plays D for the first time in some period t. Play 2 receives at least $-R(z,\delta) - (1-\delta)/\delta$ as a continuation payoff after period t, by Lemma 1 and by our assumption that R is non-increasing. Also, she receives zero in each period up to period t, because she plays L and player 1 plays U. In period t she receives $-a(1-\delta) \geq -(1-\delta)$, because she plays L and player 1 plays D and because $a \in (0,1]$. If player 1 plays U in every period, then player 2 receives zero. Player 1 will play D in some period t, with probability at most $1-z$. Consequently, player 2's payoff, if she uses the alternative strategy, is at least $-(1-z)(\delta^t(1-\delta) + \delta^{t+1}(R(z,\delta) + (1-\delta)/\delta)) \geq -(1-z)(R(z,\delta) + 2(1-\delta))$. Lemma 3, that we state below, establishes a lower bound that formalizes this line of reasoning.

The payoff that player 2 gets from the equilibrium strategy σ_2 must be least as great as the payoff she receives from the alternative strategy. So, $-(1-z)(R(z,\delta)+2(1-\delta)) \leq (1-z)(R(z,\delta)+2(1-\delta)) - zlR(z,\delta)$. Rearranging, $R(z,\delta) \leq 4(1-z)(1-\delta)/(lz-2(1-z)) \leq 4(1-\delta)/(lz-2(1-z))$. Thus, for z sufficiently close to one $R(z,\delta) \leq 8(1-\delta)/lz$.[15] So, the resistance at reputation level z is very close to zero, if δ is close to one. The argument for Corollary 1 then proceeds to show that we can redo this exercise for z' sufficiently close to z, and then z'' sufficiently close to z', working down to any $\underline{z} > 0$ that we wish, in finitely many steps.

We now proceed to establish an upper and a lower bound (Lemmata 2 and 3) for player 2's PBE payoffs. The following definition introduces a stopping time, $T(\sigma, z, z')$, which we use to construct the upper and the lower bound. $T(\sigma, z, z')$ is the first period in which player 1's reputation level exceeds z', if his initial reputation level is z and the players use strategy profile σ. For an integer T, let $E_{[0,T]}$ denote the event (set of infinite public histories) where player 1 plays D for the first time in period t for some $t \in \{0,...,T\}$. Then, $\Pr_{(\sigma_1,\sigma_2)}[E_{[0,T]}]$ is the probability that player 1 plays D, for the first time in period t for some $t \in \{0,...,T\}$, if player 1 is using strategy σ_1 and player 2 is using strategy σ_2.

Definition 2 (Stopping time)**.** *For any strategy profile $\sigma = (\sigma_1(N), \sigma_1(S), \sigma_2)$ where σ_2 is a pure strategy, reputation levels $z > 0$ and $z' > z$ let*

$$T(\sigma, z, z') = \min\{k \in \{0,1,2,...\} : z/(1-q(k)) \geq z'\},$$

[15] To be precise, if z is close to one, then $lz/2 \geq 2(1-z)$. So, $lz - 2(1-z) \geq lz/2$ and hence $R(z,\delta) \leq 4(1-\delta)/lz/2 = 8(1-\delta)/lz$.

where $q(k) = (1-z)\Pr_{(\sigma_1(N),\sigma_2)}[E_{[0,k]}]$*; and let* $T(\sigma, z, z') = \infty$ *if the set is empty.*[16]

Suppose that player 1's initial reputation level is z, σ is a PBE strategy profile and $T(\sigma, z, z')$ is the stopping time defined above. If $h^{T(\sigma,z,z')+1}$ is a history consistent with $\sigma_1(S)$ and σ_2, i.e., player 1 has always played U in all periods up to and including period T, then by definition, $\mu(S|h^{T(\sigma,z,z')+1}) \geq z'$. Also, by definition, the total probability that player 1 plays D for the first time in any period $t \in \{0, 1, ..., T(\sigma, z, z') - 1\}$ is at most $1 - z/z'$.

Lemma 2 (Upper-bound)**.** *Suppose* $0 \leq z < z' \leq 1$*. Let* $\sigma = (\sigma_1(N), \sigma_1(S), \sigma_2)$ *denote a PBE of* $\Gamma^\infty(z, \delta)$ *where player 2's resistance is at least* $R(z, \delta) - \epsilon$ *and* $\epsilon > 0$*. Then,*

$$U_2(\sigma, \delta) \leq q(R(z, \delta) + 2(1-\delta)) + R(z', \delta) + 2(1-\delta) - zl(R(z, \delta) - \epsilon) \tag{4}$$

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

Proof. Let σ_2^* denote a pure strategy in the support of σ_2 such that the resistance of σ_2^* is at least $R(z, \delta) - \epsilon$. 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 $R(z, \delta) - \epsilon$. Let profile $\sigma^* = (\sigma_1(N), \sigma_1(S), \sigma_2^*)$ and let $T = T(\sigma^*, z, z')$. By the definition of the stopping time in Definition 2, player 1's reputation exceeds z' at the end of period T, if U is played in all periods up to and including T, and if player 2 is playing according to σ_2^*. Again by Definition 2, the total probability that player 1 plays D for the first time in a period $t \in \{0, ..., T-1\}$ is at most $q = 1 - z/z'$. We bound player 2's payoffs in the following three events: (i) The event that player 1 plays D for the first time in some period $t < T$. The probability of this event is at most q. (ii) The event that player 1 plays D for the first time in some period $t \geq T$. The probability of this event is at most 1. (iii) The event that player 1 never plays D. The probability of this event is at least z, because S never plays D. These three events are exhaustive.

In a period where player 1 plays U player 2 receives at most zero. Consequently, player 2's total payoff in all the periods until player 1 plays D for the first time is at most zero. If event (i) occurs

[16]We restrict the definition above to pure strategies for player 2. For a mixed strategy for player 2, the period in which player 1's reputation exceeds z' may depend on the realization of player 2's mixture. So, if σ_2 is not a pure strategy, then the stopping time $T(\sigma, z, z')$ is a random variable. This introduces additional notation, and taking σ_2 as a pure strategy suffices for our purposes.

and player 1 plays D for the first time in period t, then player 2 receives zero until period t, receives at most $(1-\delta)$ in period t,[17] and receives a continuation payoff of at most $R(z,\delta)+(1-\delta)/\delta$ by Lemma 1, and the assumption that R is non-increasing. So, if event (i) occurs, then player 2's payoff is at most $R(z,\delta)+2(1-\delta)$ because

$$R(z,\delta)+2(1-\delta) \geq \delta^t(1-\delta)+\delta^{t+1}(R(z,\delta)+(1-\delta)/\delta) \text{ for any } t.$$

If event (ii) occurs and player 1 plays D for the first time in period t, then player 2 receives zero until period t, receives at most $(1-\delta)$ in period t, and receives a continuation payoff of at most $R(z',\delta)+(1-\delta)/\delta$, by Lemma 1, and by R non-increasing. So, if event (ii) occurs, then player 2's payoff is at most $R(z',\delta)+2(1-\delta)$ because

$$R(z',\delta)+2(1-\delta) \geq \delta^t(1-\delta)+\delta^{t+1}(R(z',\delta)+(1-\delta)/\delta) \text{ for any } t.$$

If event (iii) occurs, then player 1 plays U in each period. Player 2's payoff in this event is at most $-l(R(z,\delta)-\epsilon)$, by the definition of resistance. Putting the bounds on player 2's payoffs in the three events together implies that,

$$U_2(\sigma,\delta) \leq q(R(z,\delta)+2(1-\delta))+R(z',\delta)+2(1-\delta)-zl(R(z,\delta)-\epsilon).$$

□

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

$$U_2(\sigma,\delta) \geq -q(R(z,\delta)+2(1-\delta))-R(z',\delta)-2(1-\delta) \tag{5}$$

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

Proof. Pick any PBE σ of $\Gamma^\infty(z,\delta)$. 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 $T = T(\sigma^*, z, z')$. We again look at the following three events: (i) The event that player 1 plays D for the first time in some period $t < T$. The probability of this event is at most q. (ii) The

[17] Player 2's highest stage game payoff is one in this game because $b \leq 1$.

event that player 1 plays D for the first time in some period $t \geq T$. The probability of this event is at most 1. (iii) The event that player 1 never plays D.

Player 2's payoff until player 1 plays D for the first time is zero by definition. If event (i) occurs and player 1 plays D for the first time in period t, then player 2 receives zero until period t, receives at worst $-a(1-\delta) \geq -(1-\delta)$ in period t, and receives a continuation payoff of at worst $-R(z,\delta)-(1-\delta)/\delta$, by Lemma 1 and R non-increasing. Consequently, player 2's payoff is at least

$$-\delta^t(1-\delta) - \delta^{t+1}(R(z,\delta)+(1-\delta)/\delta) \geq -R(z,\delta) - 2(1-\delta).$$

If event (ii) occurs and player 1 plays D for the first time in period t, then player 2 receives zero until period t, receives at worst $-(1-\delta)$ in period t, and receives a continuation payoff of at worst $-R(z',\delta)-(1-\delta)/\delta$, by Lemma 1 and R non-increasing. Consequently, player 2's payoff is at least

$$-\delta^t(1-\delta) - \delta^{t+1}(R(z',\delta)+(1-\delta)/\delta) \geq -R(z',\delta) - 2(1-\delta).$$

If event (iii) occurs, then player 1 never plays D and consequently player 2 receives zero. Putting the bounds on player 2's payoffs in the three events together implies that,

$$U_2(\sigma,\delta) \geq U_2(\sigma^*,\delta) \geq -q(R(z,\delta)+2(1-\delta)) - R(z',\delta) - 2(1-\delta).$$

□

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 $z < z' \leq 1$*

$$R(z,\delta)(\underline{z}l - 2q) \leq 2R(z',\delta) + 8(1-\delta) \tag{6}$$

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

Proof. For any $\epsilon > 0$ there exists a PBE σ where player 2's resistance is at least $R(z,\delta)-\epsilon$, by the definition of resistance. By Lemma 2, equation (4) holds for any $\epsilon > 0$ and any PBE σ where player 2's resistance is at least $R(z,\delta)-\epsilon$. Also, the upper-bound in equation (4) must exceed the

lower-bound in (5) for any PBE σ. Combining (4) and (5), taking $\epsilon \to 0$, and substituting $\underline{z}$ for z implies that $R(z,\delta)(\underline{z}l - 2q) \leq 4R(z',\delta) + 4(1+q)(1-\delta)$. Using $q \leq 1$ delivers inequality (6). □

Proof of Corollary 1 under the assumption that R is non-increasing. Let $\underline{q} = \underline{z}l/4$. For any $z \in [\underline{z}, 1]$ and any $z' \in (z, 1]$ such that $1 - z/z' \leq \underline{q}$ inequality (6) implies

$$R(z,\delta)(\underline{z}l - 2\underline{q}) \leq 2R(z',\delta) + 8(1-\delta) \quad (\text{using } q = 1 - z/z' \leq \underline{q})$$

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

$$R(z'(1-q),\delta) \leq \frac{4}{\underline{z}l}(R(z',\delta) + 4(1-\delta)). \quad (\text{substituting } z'(1-q) \text{ for } z) \tag{7}$$

Notice $R(1,\delta) = 0$. For $z = 1 - \underline{q}$ and $z' = 1$ inequality (7) and $R(1,\delta) = 0$ imply that $R(1-\underline{q},\delta) \leq 16(1-\delta)/\underline{z}l$. Again, for $z = (1-\underline{q})^2$ and $z' = 1 - \underline{q}$, inequality (7) and $R(1-\underline{q},\delta) \leq 16(1-\delta)/\underline{z}l$ imply that $R((1-\underline{q})^2,\delta) \leq 64(1-\delta)/(\underline{z}l)^2 + 16(1-\delta)/\underline{z}l$. More generally, for any $z \geq \underline{z}$,

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

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

3.1. **Description of the proof of Theorem 1.** Our discussion up to this point established a reputation result for the example depicted in Figure 5. However, under the assumptions of $\mu(\Omega_-) = 0$ and R non-increasing, the same argument works, with minor modifications, for any stage game of perfect information that satisfies Assumption 1 and $n^p = 1$. We now sketch the steps involved in allowing for $n^p > 1$, $\mu(\Omega_-) > 0$ and relaxing the assumption that R is non-increasing.

Lemmata A.1 and A.2 are the technical steps that allow us to accommodate a more complicated dynamic Stackelberg type who may punish player 2, i.e., the case where $n^p > 1$. Lemma A.1 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 phase if he instead plays according to $\sigma_1(S)$ in order to maintain his reputation. Lemma A.2 is an analog of Lemma 1 that accounts for these punishment phases.

Allowing for other commitment types, i.e., $\mu(\Omega_-) > 0$, 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 $\mu(\Omega_-)/\mu(S)$ is non-increasing if player 1 plays according to $\sigma_1(S)$. This monotonicity allows us to treat $\mu(\Omega_-)/\mu(S)$ as an additional state variable in Definitions A.2 and A.3. 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 another 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 two reputation levels $z \in [\underline{z},1]$ 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 equation (6). Rewriting,

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

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 $\underline{z}D_1 - D_2/K - D_3/K \geq \underline{z}D_1/2$ (equation (10)). Let $z_n(\delta)$ be the supremum over reputation levels z such that $R(z,\delta) \geq K^n(1-\delta)$ (Definition A.2). 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 $q_n(\delta) = 1 - z_n(\delta)/z_{n-1}(\delta)$. 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, $\lim_{\delta\to 1} R(z,\delta) = 0$.

4. Discussion

4.1. **Necessity of perfect information.** Perfect information is necessary for a reputation result in repeated locally non-conflicting interests games. Without perfect information, a folk theorem applies for example to the simultaneous-move common interest game in Figure 3a (Cripps and Thomas (1997)), which is a locally non-conflicting interest game. For a reputation result in repeated strictly conflicting interests stage games, perfect information assumption is not required (see Cripps et al. (2005) or section 4.5).

Figure 5 is a normalized sequential common interest game if $a = 1$ and $b = -1$. Consequently, Corollary 1 is a particular example of a reputation result for a repeated locally non-conflicting interests game (the sequential common interest game). Lemma 1 is central for establishing Corollary 1 and the perfect information assumption is required for Lemma 1. In order to flesh out the intuition of why perfect information is necessary, we construct a PBE for the repeated simultaneous move common interest game given in Figures 3a, where there is no analog of Lemma 1. In this PBE, the players' payoffs are low, if z is close to zero and δ is close to one.[18] That is, the failure of Lemma 1 also leads to the failure of the reputation result.

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, the continuation payoff for the players, 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 L and player 1 plays D (thus revealing rationality). Punishment entails a continuation payoff for player 2 that is close to zero.[19] 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 (L,D), she can be deterred from playing L even if player 1 reveals rationality with a small probability in each period.

[18] This construction follows Cripps and Thomas (1997).

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

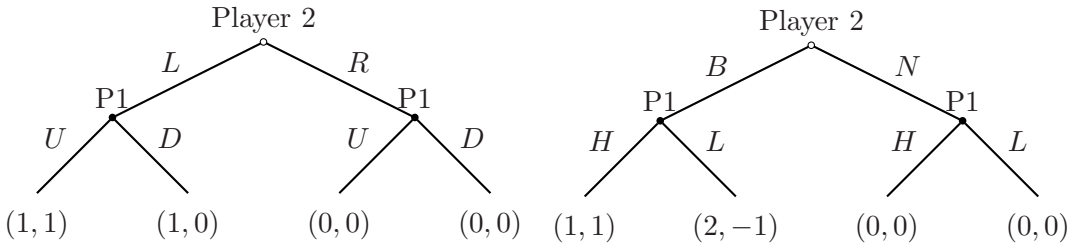
However, if the probability that player 1 reveals rationality is small in each period, then it takes many periods for player 1 to build a reputation and K can be chosen large to ensure low payoffs for both players.

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 thus we can put no restrictions on payoffs at* (L, D). In contrast, if player 1 moves after observing player 2 as in Figures 3b, then Lemma 1 implies that player 2's continuation payoff after (L, D) is at least $-2R(z/(1-q), \delta) - 2(1-\delta)$, i.e., Lemma 1 imposes a tight bound on the amount of punishment that player 2 can expect after choosing L.

For our reputation result we make extensive use of Lemma 1 in establishing the upper and lower bounds for player 2's payoffs (Lemma 2 and Lemma 3). 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. The bound in Lemma 1 applies verbatim to the simultaneous-move game at node (R, D) (which is the node of interest for Lemma 2), because player 1 believes that player 2 plays R with probability one on the equilibrium path.

In contrast to Lemma 2, perfect information is essential for Lemma 3. 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). Lemma 1 provides a lower-bound on player 2's payoff after (L, D) in the case of perfect information. However, there is no analog of Lemma 1 that provides a tight bound on player 2's payoff after (L, D) for the simultaneous-move game. For example, in the PBE we construct we can put no restrictions on payoffs after node (L, D) beyond individual rationality and feasibility. This is because player 1 expects to reach node (L, D) with probability zero.

4.2. **Necessity of Assumption 1.** Assumption 1 can fail in two ways. First, Assumption 1 fails if the payoff profile where player 1 receives $\bar{g}_1$ is not unique in G, for example if Γ is non-generic. Such a failure is depicted in Figure 7a. Second, Assumption 1 fails if, $(\bar{g}_1, \hat{g}_2) \in G$, but Γ is not a strictly conflicting interests game. Such a failure is depicted in Figure 7b. A reputation result also fails to obtain in both of these examples.



(A) A common interest game that is non-generic.

(B) A moral hazard mixing game.

FIGURE 7. Games that fail to satisfy Assumption 1.

In the non-generic common interest game depicted in Figure 7a 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 7b 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: Suppose that 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.[20] 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, $\mu(S)\epsilon < \epsilon/2$.

4.3. **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 4, 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.

4.4. **Other commitment types.** As noted previously by Schmidt (1993), Celantani et al. (1996) or Evans and Thomas (1997), if there a chance that player 1 is a commitment type, other than the Stackelberg type, then player 1 may be unable 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 can also hinder player 1 from building a reputation. A patient player 2 may resist the Stackelberg type because she fears punishment 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 because we assume that the probability of another commitment type is small compared to the probability of the Stackelberg type.

The 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

[20]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.

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.[21] 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.[22]

The reputation result of Schmidt (1993) obtains if there are conflicting interests in the stage game, 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 strictly conflicting interests.

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

[22] See Atakan and Ekmekci (2008) which assumes player 2's moves are imperfectly 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.

4.5. **Games with strictly conflicting interests.** Cripps, Dekel, and Pesendorfer (2005) obtain a reputation result for Bayes-Nash equilibria and simultaneous-move strictly conflicting interests 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 A.3 remains valid for Bayes-Nash equilibria. This is because all the arguments were constructed on the 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, a functional inequality similar to (6) holds, and a reputation result follows.

Appendix A. Proof of Theorem 1

Let $M = \max\{\max\{|g_1|,|g_2|\} : (g_1,g_2) \in F\}$ and normalize payoffs, without loss of generality, such that

$$\bar{g}_1 = 1;\ g_1(a_1,a_2) \geq 0 \text{ for all } a \in A;\ \text{and } g_2(a_1^s,a_2^b) = 0. \tag{8}$$

For any $z \in (0,1)$ let

$$K(z) = \max\left\{\frac{4\rho(1+2n^pM)+16M}{lz}, 1\right\}.$$

For any $z \in (0,1)$ let

$$f(z) = \frac{K(z)^{\bar{n}(z)}}{z} \tag{9}$$

where $\bar{n}(z)$ is the smallest integer j such that $(1-zl/4\rho)^j < z$. In what follows fix $\underline{z} > 0$, fix

$$K = K(\underline{z}),\ \text{and } \bar{n} = \bar{n}(\underline{z}). \tag{10}$$

We show that

$$U_1(\sigma,\delta) \geq 1 - f(\underline{z})\max\{1-\delta,\mu(\Omega_-)\} = 1 - \frac{K^{\bar{n}}}{\underline{z}}\max\{1-\delta,\mu(\Omega_-)\}$$

for any $\mu \in \Delta(\Omega)$ such that $\mu(S) = \underline{z}$ and any PBE strategy profile σ of $\Gamma^\infty(\mu,\delta)$.

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

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

where $r(\sigma_2,\delta)=1-U_1(\sigma_1(S),\sigma_2,\delta)$ *for any* σ_2.

Lemma A.1. *Posit perfect information and Assumption 1. There exists* $\delta^*\in[0,1)$ *and* $l>0$ *such that for any* $r\geq 0$, *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 (3) implies that there exists a $\delta^*<1$ and $l>0$ such that, for all $\delta>\delta^*$,

$$g_2(a_1^s,a_2)+\sum_{k=1}^{n^p-1}\delta^k g_2(a_1^p,a_2')<-ln^p \tag{11}$$

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 $(1-\delta)\mathbb{E}_{(\sigma_1(S),\sigma_2)}\left[\sum_{t=0}^{\infty}\delta^t i(h^t)\right]\geq r/n^p$.[23] 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 (11). 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* $\delta>\delta^*$. Also, *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.

Lemma A.2. *Pick any PBE* σ *of* $\Gamma^\infty(\mu,\delta)$, *period* t *public history* $h=(h^t,d_0)$, *and suppose player 1 is to deviate from* $\sigma_1(S)$ *at node* d_0 *with positive probability given* h. *Let* h^{t+1} *be any public history of terminal nodes that is reached with positive probability under* $\Pr_{(\sigma_1(N)|_h,\sigma_2|_h)}$; *let* $h'=(h^t,d')$ *be the public history that is reached immediately (with positive probability under* $\Pr_{(\sigma_1(S)|_h,\sigma_2|_h)}$*) if* $\sigma_1(S)$ *is used at* d *instead of deviating; and let* $\mu'=\mu(\cdot|h')$, *then*

$$|U_2(\sigma_1(N),\sigma_2,\delta|h^{t+1})|\leq \rho(R(\mu',\delta)+n^pM(1-\delta)/\delta),\ \textit{if}\ \Gamma\ \textit{satisfies Ass. 1 (i), and}$$

$$U_2(\sigma_1(N),\sigma_2,\delta|h^{t+1})\leq \rho(R(\mu',\delta)+n^pM(1-\delta)/\delta),\ \textit{if}\ \Gamma\ \textit{satisfies Ass. 1 (ii)}.$$

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

Proof. If player 1 plays according to $\sigma_1(S)$ at d_0 and through the remaining nodes of period t, then he obtains at lest zero for the period and an $n^p - 1$ period punishment phase may ensue. His payoff is at least zero in these periods. So, his payoff if he plays according to $\sigma_1(S)$ is at least $\delta^{n^p}(1 - R(\mu', \delta))$. Alternatively, if he plays according to $\sigma_1(N)$ and deviates from $\sigma_1(S)$, he receives at most $M(1-\delta)$ for the period, and $U_1(\sigma, \delta|h^{t+1})$ as his continuation payoff. So, $M(1-\delta) + \delta U_1(\sigma, \delta|h^{t+1}) \geq \delta^{n^p}(1 - R(\mu', \delta))$. This implies that

$$U_1(\sigma, \delta|h^{t+1}) \geq \delta^{n^p-1}(1 - R(\mu', \delta)) - M(1-\delta)/\delta \geq 1 - R(\mu', \delta) - n^p M(1-\delta)/\delta.$$

The bounds on player 2's payoff follow from equations (1), (2) and $(U_1(\sigma, \delta|h^{t+1}), U_2(\sigma_1(N), \sigma_2, \delta|h^{t+1})) \in F$. □

Pick any period t public history $h = (h^t, d_0)$, 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 A.4 by considering a non-equilibrium strategy for player 2 and applying Lemma A.2.

Definition A.2 (Reputation Thresholds)**.** *For each $n \geq 0$, let*

$$z_n(\delta, \phi) = \sup\{z : \exists \mu \in \Delta(\Omega) \text{ s.t. } R(\mu, \delta) \geq K^n \max\{\phi, 1-\delta\}, \mu(S) = z, \ \mu(\Omega_-)/\mu(S) \leq \phi\},$$

where K is the constant defined in equation (10).

Definition A.3. *For any $\xi > 0$ and $z \in (0,1)$ let*

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

By definition, there exists μ such that $\mu(S) = z \in [z_n(\delta, \phi) - \xi, z_n(\delta, \phi)]$ and $\mu(\Omega_-)/\mu(S) \leq \phi$, 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, $\bar{R}(\xi, z_n, \delta, \phi) \geq K^n \max\{\phi, 1-\delta\}$. The definition of $z_n(\delta, \phi)$ and $\bar{R}(\xi, z_n, \delta, \phi) \geq K^n \max\{\phi, 1-\delta\}$ implies that if $\mu(S) \in [z_n(\delta, \phi) - \xi, z_{n-1}(\delta, \phi)]$ and $\mu(\Omega_-)/\mu(S) \leq \phi$, then $R(\mu, \delta) \leq \bar{R}(\xi, z_n, \delta, \phi)$ in any PBE profile σ of $\Gamma^\infty(\mu, \delta)$. In what follows we establish a upper bound on Player 2's payoff

in any PBE where the resistance is at least $\bar{R}(\xi, z_n, \delta, \phi) - \xi$ and a lower bound on player 2's PBE payoff. First we introduce a stopping time that we use for the argument.

Definition A.4 (Stopping time)**.** *For any 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 $0 \leq t \leq T$. 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 $z' \in (\mu(S), 1]$ let*

$$T(\sigma, \mu, z') = \min\{t : \mu(S)/(1 - q(t)) \geq z'\},$$

where $q(t) = \sum_{\omega\in\Omega} \mu(\omega) \Pr_{(\sigma_1(\omega),\sigma_2)} \left[E_{[0,T]}\right]$; and let $T(\sigma, \mu, q) = \infty$ if the set is empty.

Suppose that player 1's initial reputation level $\mu(S) = z$ and $\mu(\Omega_-)/\mu(S) \leq \phi$. 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, by definition, then, $\mu(S|h^{T(\sigma,z,z')+1}) \geq z'$. Also, by definition, the total probability that player 1 deviates from the Stackelberg strategy for the first time in any period $t \in \{0, 1, ..., T(\sigma, z, z') - 1\}$ is at most $1 - z/z'$, and Bayes' rule implies that $\mu(\Omega_-|h^{T(\sigma,\mu,q)+1})/\mu(S|h^{T(\sigma,\mu,q)+1}) \leq \phi$.

Lemma A.3. *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 $\mu(\Omega_-)/\mu(S) \leq \phi$, 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,*

(12)

$$U_2(\sigma, \delta) \leq \rho(q(\delta, \phi, n, \xi)\bar{R}(\xi, z_n, \delta, \phi) + K^{n-1}\epsilon + 2n^p M\epsilon) + 5M\epsilon - (\bar{R}(\xi, z_n, \delta, \phi) - \xi)(z_n(\delta, \phi) - \xi)l,$$

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

Proof. Choose pure strategy σ_2^* in the support of the possibly mixed strategy σ_2 such that $r(\sigma_2^*, \delta) \geq \bar{R}(\xi, z_n, \delta, \phi) - \xi$. Since the mixed strategy has resistance equal to $\bar{R}(\xi, z_n, \delta, \phi) - \xi$, there must be a pure strategy in the support of this mixed strategy which has resistance of at least $\bar{R}(\xi, z_n, \delta, \phi) - \xi$. 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 $\mu(\Omega_-)/\mu(S) \leq \phi$ if player 1 has not deviated from $\sigma_1(S)$ in h^t that is consistent with σ_2^*, then $\mu(\Omega_-|h^t)/\mu(S|h^t) \leq \phi$; 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 $E_{[0,T-1]}$; $\omega = N$ and $E_{[T,\infty)}$; $\omega = N$ and he never deviates from $\sigma_1(S)$; $\omega = S$; 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 A.1 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 $h^\infty \in E_{[0,T-1]}$ and let $h = (h^j, d)$ denote the node in period j where player 1 deviates from $\sigma_1(S)$ for the first time in the infinite public history h^∞. Player 2's payoff up until period j is at most ϵM. Player 2's payoff in period j is at most ϵM. Lemma A.2 and $\epsilon \geq (1-\delta)$ imply that $U_2(\sigma_1(N), \sigma_2, \delta | h^{j+1}) \leq \rho(\bar{R}(\xi, z_n, \delta, \phi) + \epsilon M n^p)$. Hence, for any such period j player 2's payoff is at most

$$M\epsilon + \delta^j M\epsilon + \delta^{j+1}\rho(\bar{R}(\xi, z_n, \delta, \phi) + \epsilon M n^p) \leq 2M\epsilon + \rho(\bar{R}(\xi, z_n, \delta, \phi) + n^p M\epsilon)$$

So,

$$U_2(\sigma_1(N), \sigma_2, \delta | E_{[0,T-1]}) \leq 2M\epsilon + \rho(\bar{R}(\xi, z_n, \delta, \phi) + n^p M\epsilon). \tag{13}$$

Suppose that $h^\infty \in E_{[T,\infty)}$ and let $h = (h^j, d)$ denote the node where player 1 deviates from $\sigma_1(S)$ for the first time in the infinite public history h^∞. Player 1's reputation exceeds z_{n-1} at the start of period $j+1$ if he plays according $\sigma_1(S)$ through period j. Consequently, resistance is at most $K^{n-1}\epsilon$ at the start of period $j+1$ and Lemma A.2 implies that $U_2(\sigma_1(S), \sigma_2, \delta | h^{j+1}) \leq \rho\epsilon(K^{n-1} + n^p M)$. So, an argument identical to the previous paragraph implies that

$$U_2(\sigma_1(N), \sigma_2, \delta | E_{[T,\infty)}) \leq 2M\epsilon + \rho\epsilon(K^{n-1} + n^p M). \tag{14}$$

If player 1 never deviates from $\sigma_1(S)$, then player 2 receives at most zero. 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 A.1, and this happens with probability $z \geq z_n(\delta, \phi) - \xi$. The probability of N and $E_{[0,T-1]}$ is at most $q(\delta, \phi, n, \xi)$; and the probability of N and $E_{[T,\infty)}$ is at most one. Consequently, equations (13) and (14) imply

$$U_2(\sigma, \delta) \leq q(\delta, \phi, n, \xi)\rho\bar{R}(\xi, z_n, \delta, \phi) + \rho K^{n-1}\epsilon - (z_n(\delta, \phi) - \xi)(\bar{R}(\xi, z_n, \delta, \phi) - \xi)l + 2\rho n^p M\epsilon + 5M\epsilon.$$

□

Lemma A.4. *Posit perfect information and Assumption 1 item (i). Suppose that* $\mu(S) = z \in [z_n(\delta,\phi) - \xi, z_n(\delta,\phi)]$ *and* $\mu(\Omega_-)/\mu(S) \leq \phi$. *In any PBE* σ *of* $\Gamma^\infty(\mu,\delta)$,

$$U_2(\sigma,\delta) \geq -\rho(\bar{R}(\xi,z_n,\delta,\phi)q(\delta,\phi,n,\xi) + K^{n-1}\epsilon + 2n^pM\epsilon) - 3M\epsilon, \tag{15}$$

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

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 $\mu(\Omega_-)/\mu(S) \leq \phi$. 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 $(\sigma_1(S), \sigma_2^*)$. Also, player 2's payoff in the period player 1 deviates from $\sigma_1(S)$ is at least $-M\epsilon$. We use the reasoning in Lemma A.3 and apply Lemma A.2 to obtain $U_2(\sigma_1(N), \sigma_2, \delta|E_{[0,T-1]}) \geq -\rho(\bar{R}(\xi,z_n,\delta,\phi) + n^pM\epsilon) - M\epsilon$ and $U_2(\sigma_1(N), \sigma_2, \delta|E_{[T,\infty)}) \geq -\rho\epsilon(K^{n-1} + n^pM) - M\epsilon$. If player 1 never deviates from $\sigma_1(S)$, then player 2 receives zero. 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 same reasoning as in Lemma A.3 implies that

$$U_2(\sigma,\delta) \geq U_2(\sigma^*,\delta) \geq -\rho\bar{R}(\xi,z_n,\delta,\phi)q(\delta,\phi,n,\xi) - \rho\epsilon K^{n-1} - 2\rho n^pM\epsilon - 3M\epsilon.$$

□

Completing the argument for Theorem 1 by using Lemma A.3 and Lemma A.4. If Γ satisfies Assumption 1 and perfect information, then equation (12) is satisfied, by Lemma A.3. If Γ satisfies Assumption 1 (i) and perfect information, then equation (15) is satisfied, by Lemma A.4. Also, if Γ satisfies Assumption 1 item (ii), then $U_2(\sigma,\delta) \geq \hat{g}_2 = 0$, and equation (15) is trivially satisfied because the right hand side of the inequality is negative. Combining the upper and lower bounds for $U_2(\sigma,\delta)$, given by equations (12) and (15), and simplifying by canceling ϵ delivers

$$(z_n(\delta,\phi) - \xi)l\frac{\bar{R}(\xi,z_n,\delta,\phi) - \xi}{\epsilon} \leq 2\rho\left(\frac{q(\delta,\phi,\xi)\bar{R}(\xi,z_n,\delta,\phi)}{\epsilon} + K^{n-1} + 2n^pM\right) + 8M.$$

Let $q_n(\delta,\phi) = 1 - z_n(\delta,\phi)/z_{n-1}(\delta,\phi)$. $\bar{R}(\xi, z_n, \delta, \phi) \in [0,1]$ 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

$$z_n(\delta,\phi) l \bar{R}(z_n,\delta,\phi)/\epsilon \leq 2\rho(q_n(\delta,\phi)\bar{R}(z_n,\delta,\phi)/\epsilon + K^{n-1} + 2n^p M) + 8M.$$

Rearranging,

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

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,

$$q_n(\delta,\phi) \geq \frac{z_n(\delta,\phi)l}{2\rho} - \frac{K^{n-1}}{K^n} - \frac{2n^p M}{K^n} - \frac{4M}{\rho K^n}.$$

So, $q_n(\delta,\phi) \geq \frac{\underline{z}l}{2\rho} - \frac{1}{K} - \frac{2n^p M}{K^n} - \frac{4M}{\rho K^n}$, for any $z_n(\delta,\phi) \geq \underline{z}$. The definition of K, which is given in equation (10), implies that

$$\frac{\underline{z}l}{2\rho} - \frac{1}{K} - \frac{2n^p M}{K^n} - \frac{4M}{\rho K^n} \geq \frac{\underline{z}l}{2\rho} - \frac{1}{K} - \frac{2n^p M}{K} - \frac{4M}{\rho K} \geq \frac{\underline{z}l}{4\rho} > 0$$

for any $n \geq 1$. Consequently, $q_n(\delta,\phi) \geq \frac{\underline{z}l}{4\rho} > 0$, for any $z_n(\delta,\phi) \geq \underline{z}$. So, $z_n(\delta,\phi) \geq \underline{z}$ implies that $1 - z_n(\delta,\phi)/z_{n-1}(\delta,\phi) \geq \frac{\underline{z}l}{4\rho}$ for all $\delta < 1$, $\phi > 0$ and $n = 0, 1, ..., \infty$. Also, the definition of $\bar{n}$, which is given in equation (10), requires that $(1 - \frac{\underline{z}l}{4\rho})^{\bar{n}} < \underline{z}$. So, the definition of $\bar{n}$ implies that $z_{\bar{n}}(\delta,\phi) \leq (1 - \frac{\underline{z}l}{4\rho})^{\bar{n}} < \underline{z}$, for each $\delta < 1$ and $\phi > 0$. Consequently, if $\mu(S) \geq \underline{z}$ and $\mu(\Omega_-)/\mu(S) \leq \phi$, then $R(\mu,\delta) \leq K^{\bar{n}} \max\{1-\delta, \phi\}$ and $U_1(\sigma,\delta) \geq 1 - K^{\bar{n}} \max\{1-\delta,\phi\} \geq 1 - K^{\bar{n}}/\underline{z} \max\{1-\delta, \mu(\Omega_-)\}$ 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*.

Bar-Isaac, H. (2003): "Reputation and survival: learning in a dynamic signalling model," *The Review of Economic Studies*, 70(2), 231–251.

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 (1995): "Reputation and commitment in two-person repeated games without discounting," *Econometrica: Journal of the Econometric Society*, pp. 1401–1419.

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

Cripps, M., and J. Thomas (2003): "Some asymptotic results in discounted repeated games of one-sided incomplete information," *Mathematics of Operations Research*, pp. 433–462.

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

Faingold, E. (2005): "Building a reputation under frequent decisions," Discussion paper, mimeo.

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.

Peski, M. (2008): "Repeated games with incomplete information on one side," *Theoretical Economics*, 3(1), 29–84.

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.

Wiseman, T. (2005): "A partial folk theorem for games with unknown payoff distributions," *Econometrica*, 73(2), 629–645.

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