Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 03-016

"Imperfect Monitoring and Impermanent Reputations"

by

Martin W. Cripps, George J. Mailath and Larry Samuelson

http://ssrn.com/abstract=418600

IMPERFECT MONITORING AND IMPERMANENT REPUTATIONS[1]

By Martin W. Cripps, George J. Mailath, and Larry Samuelson

May 30, 2003
first version July 30, 2002

Abstract: We study the long-run sustainability of reputations in games with imperfect public monitoring. It is impossible to maintain a permanent reputation for playing a strategy that does *not* play an equilibrium of the game without uncertainty about types. Thus, a player cannot indefinitely sustain a reputation for non-credible behavior in the presence of imperfect monitoring.

Journal of Economic Literature Classification Numbers C70, C78.
Keywords: Reputation, Imperfect Monitoring, Repeated Games, Commitment, Stackelberg types.

[1] We thank Glenn Ellison, three referees, Drew Fudenberg, Ehud Kalai, and audiences at many seminars for their helpful comments and the National Science Foundation (grants SES-0095768 and SES-9911219) for financial support.

1 Introduction

The adverse selection approach to reputations is central to the study of long-run relationships. In the finitely-repeated prisoners' dilemma or chain-store game, for example, the intuitive expectation that cooperation or entry deterrence occurs in early rounds is inconsistent with equilibrium. However, incomplete information about a player's characteristics can be exploited to support an equilibrium reputation for cooperating or fighting entry (Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982), Milgrom and Roberts (1982)). In infinitely repeated games, the multiplicity of equilibria provided by the folk theorem contrasts with the intuitive attraction of equilibria that provide relatively high payoffs. Reputation effects can again rescue intuition by imposing lower bounds on equilibrium payoffs (Fudenberg and Levine (1989, 1992)).

This paper explores long-run reputation effects in games of imperfect monitoring with a long-lived player facing a sequence of short-lived players. In the absence of incomplete information about the long-lived player, her equilibrium payoff can be any value between her minmax payoff and an upper bound (independent of her discount factor) strictly smaller than her Stackelberg payoff. However, when there is incomplete information about the long-lived player's type, reputation effects imply that the equilibrium payoff of a patient long-lived player must be arbitrarily close to her Stackelberg payoff (Fudenberg and Levine (1992)).

This powerful implication is a "short-run" reputation effect, concerning the long-lived player's expected average payoff *calculated at the beginning of the game*. We show that this implication does *not* hold in the long run: A long-lived player can maintain a permanent reputation for playing a commitment strategy in a game with imperfect monitoring only if that strategy plays an equilibrium of the corresponding complete-information stage game.

More precisely, the long-lived player in the incomplete-information game is either a commitment type, who plays an exogenously specified stage-game action, or a normal type, who maximizes payoffs. The actions, and hence beliefs, of the uninformed short-lived players are public, so that the long-lived player's reputation is public. We show that if the commitment action is *not* an equilibrium strategy for the normal type in the stage game, then in any Nash equilibrium of the incomplete-information repeated game, almost surely the short-lived players will learn the long-lived player's type. Thus, a long-lived player cannot indefinitely maintain a reputation for behavior that is not credible given the player's type.

The assumption that monitoring is imperfect is critical.[2] It is straightforward to construct equilibria under perfect monitoring that exhibit permanent reputations. Any deviation from the commitment strategy reveals the type of the deviator and triggers a switch to an undesirable equilibrium of the resulting complete-information continuation game. In contrast, under imperfect monitoring, any deviation by the long-lived player neither reveals the deviator's type nor triggers a punishment. Instead, the long-run convergence of beliefs ensures that eventually *any* current signal of play has an arbitrarily small effect on the short-lived player's beliefs. As a result, a long-lived player ultimately incurs virtually no cost from a single small deviation from the commitment strategy. But the long-run effect of many such small deviations from the commitment strategy is to drive the equilibrium to full revelation. Reputations can thus be maintained only in the absence of an incentive to indulge in such deviations, that is, only if the reputation is for behavior that is part of an equilibrium of the complete-information stage game.

The impermanence of reputation arises at the behavioral as well as at the belief level. Asymptotically, continuation play is a Nash equilibrium of the complete-information game. Moreover, while the explicit construction of equilibria in reputation games is difficult, we are able to provide a partial converse (under a continuity hypothesis): for any strict Nash equilibrium of the stage game and $\varepsilon > 0$, there is a Nash equilibrium of the incomplete-information game such that if the long-lived player is normal, then with probability at least $1 - \varepsilon$, eventually the stage-game Nash equilibrium is played in every period.[3]

While the short-run properties of equilibria are interesting, we believe that the long-run equilibrium properties are relevant in many situations. For example, an analyst may not know the age of the relationship to which the model is to be applied. We do sometimes observe strategic interactions from a well-defined beginning, but we also often encounter on-going interactions whose beginnings are difficult to identify. Long-run equilibrium properties may be an important guide to behavior in the latter cases. Alternatively,

[2]Our results do apply to games of perfect monitoring in which the commitment type plays a mixed strategy (see the discussion at the conclusion of Section 4.1).

[3]Since these results hold for any discount factor, there is an apparent tension with Fudenberg and Levine (1992). However, the typical exercise in the reputation literature is to fix the prior probability of the commitment type, and then take the discount factor close to one. We instead fix both the prior and the discount factor (which may be close to one, given the prior), and examine long-run equilibrium behavior. The posterior probability of the commitment type eventually becomes small given the discount factor.

one might take the view of a social planner who is concerned with the continuation payoffs of the long-run player and with the fate of all short-run players, even those in the distant future. Our analysis also suggests that the short-run players may have definite preferences as to where they appear in the queue of short-run players, offering a new perspective on the incentives created by repeated games. Finally, interest often centers on the *steady states* of models with incomplete information, again directing attention to long-run properties.

We view our results as suggesting that a model of *long-run* reputations should incorporate some mechanism by which the uncertainty about types is continually replenished. For example, Holmström (1999), Cole, Dow, and English (1995), Mailath and Samuelson (2001), and Phelan (2001) assume that the type of the long-lived player is governed by a stochastic process rather than being determined once and for all at the beginning of the game. In such a situation, reputations can indeed have long-run implications.

The next section uses a simple motivating example to place our contribution in the literature. Section 3 describes our model. Section 4 presents the statements of the theorems, with the main result proven in Section 5. For expositional clarity, most of the paper considers a long-lived player, who can be one of two possible types—a commitment type who always plays the same (possibly mixed) stage-game action and a normal type—facing a sequence of short-lived players whose actions are perfectly observed. Section 6 provides conditions under which our results continue to hold when there are many possible commitment types, when these commitment types play more complicated strategies, when the uninformed player is long-lived, and when the short-run player's actions are not observed.

2 Related Literature

Consider an infinitely-lived player 1 with discount factor δ playing a simultaneous-move stage game with a succession of short-lived player 2's who each live for one period. The stage game is given by

(1)

1 \ 2	L	R
T	$2,3$	$0,2$
B	$3,0$	$1,1$

and has a unique Nash equilibrium, BR, which is strict.

Player 1's action in any period is not observed by any player 2. There is, however, a public signal of player 1's action, that takes on two possible values, y' and y'', according to the distribution

$$\Pr\{\ y = y' \mid i\ \} = \begin{cases} p, & \text{if } i = T, \\ q, & \text{if } i = B, \end{cases}$$

where $p > q$. Player 2's actions are public. Player 1's payoffs are as in the above stage game (1), and player 2's ex post payoffs (i.e., payoffs as a function of the realized public signal and his own action) are given by

	L	R
y'	$3(1-q)/(p-q)$	$(1-2q+p)/(p-q)$
y''	$-3q/(p-q)$	$(-2q+p)/(p-q)$

Expected payoffs for player 2 are thus still given by (1). This structure of ex post payoffs ensures that the information content of the public signal is identical to that of player 2's payoffs.

This game is an example of what Fudenberg and Levine (1994) call a *moral hazard mixing game*. Even for large δ, the long-run player's maximum Nash (or, equivalently, sequential) equilibrium payoff is lower than when monitoring is perfect (Fudenberg and Levine (1994, Theorem 6.1, part (iii))).[4] For our example, it is straightforward to apply the methodology of Abreu, Pearce, and Stacchetti (1990) to show that if $2p > 1 + q$, the set of Nash equilibrium payoffs for large δ is given by the interval

$$\left[1,\ 2 - \frac{(1-p)}{(p-q)}\right]. \tag{2}$$

Moreover, if $2\delta(p-q) > 1$, there is a continuum of particularly simple equilibria, with player 1 placing equal probability on T and on B in every period, irrespective of history, and with player 2's strategy having one period memory. Player 2 plays L with probability α' after signal y' and with probability α'' after signal y'', with

$$2\delta(p-q)\left(\alpha' - \alpha''\right) = 1.$$

The maximum payoff of $2 - (1-p)/(p-q)$ is obtained by setting $\alpha' = 1$.

We introduce incomplete information by assuming there is a probability $p_0 > 0$ that player 1 is the Stackelberg type who plays T in every period.

[4] In other words, the folk theorem of Fudenberg, Levine, and Maskin (1994) does not hold when there are short-lived players.

Fudenberg and Levine (1992) show that for *any* payoff $u < 2$, there is δ sufficiently close to 1 such that in every Nash equilibrium, the expected average discounted payoff to player 1 is at least u. We emphasize that u can exceed the upper bound in (2), so that the normal player 2 does strictly better in every equilibrium of the incomplete-information game than the best complete-information equilibrium.[5]

Our result is that the effect of the incomplete information about player 1, including the lower bounds placed on payoffs illustrated in this example, is temporary. To develop intuition, consider a Markov perfect equilibrium, with player 2's belief that player 1 is the Stackelberg type (i.e., player 1's "reputation") being the natural state variable. In any such equilibrium, the normal type cannot play T for sure in any period: if she did, the posterior after any signal in that period equals the prior, and hence continuation play is independent of the signal. But then player 1 has no incentive to play T. Thus, in any period of a Markov perfect equilibrium, player 1 must put positive probability on B. Consequently, the signals are continually informative about player 1's type, and so almost surely, when player 1 is normal, beliefs converge to zero probability on the Stackelberg type.[6] Our analysis exploits this intuition, but we do not restrict attention to Markov perfect equilibria and we generalize the result to more complicated commitment types.

While some of our arguments and results are reminiscent of the recent literature on rational learning and merging, there are important differences. For example, Jordan (1991) studies the asymptotic behavior of "Bayesian strategy processes," in which myopic players play a Bayes-Nash equilibrium of the one-shot game in each period, players initially do not know the payoffs of their opponents, and players observe past play. The central result is that play converges to a Nash equilibrium of the complete-information stage game. In contrast, the player with private information in our game is long-lived and potentially very patient, introducing intertemporal considerations that do not appear in Jordan's model, while the information processing in our model is complicated by the imperfect monitoring.

A key idea in our results (in particular, Lemma 1) is that if signals are statistically informative about a player's behavior, then nontrivial beliefs about that players's type can persist only if different types asymptotically

[5] For any $u < 5/2$, if the commitment type is a mixed commitment type, playing T with a probability less than but sufficiently close to $1/2$ and δ is sufficiently close to one, then every Nash equilibrium average discounted payoff for player 1 must be at least u.

[6] Benabou and Laroque (1992) study the Markov perfect equilibrium of a game with similar properties. They show that player 1 eventually reveals her type in any Markov perfect equilibrium.

play identically. Similar ideas play an important role in merging arguments (e.g., Sorin (1999)), which provide conditions under which a stochastic process and beliefs over that process converge. Kalai and Lehrer (1995) use merging to provide a simple argument that in reputation games, asymptotic continuation play is a subjective correlated equilibrium of the complete-information game. This result is immediate in our context, since we begin with a Nash equilibrium of the incomplete-information game (in contrast to the weaker assumptions of Kalai and Lehrer (1995)).

Jackson and Kalai (1999) prove that if a finitely repeated normal-form game with incomplete information (for which Fudenberg and Maskin (1986) prove a reputation folk theorem) is itself repeated, with new players in each repetition, then eventually, reputations cannot affect play in the finitely repeated game. While they reach a similar conclusion, the model is quite different. In particular, players in one round of the finitely repeated game do not internalize the effects of their behavior on beliefs and so behavior of players in future rounds, and there is perfect monitoring of actions in each stage game. We exploit the imperfection of the monitoring to show that reputations are eventually dissipated even when players recognize their long-run incentives to preserve these reputations.

3 The Model

3.1 The Complete-Information Game

The stage game is a two-player simultaneous-move finite game of public monitoring. Player 1 chooses an action $i \in \{1, 2, \ldots, I\} \equiv I$ and player 2 simultaneously chooses an action $j \in \{1, 2, \ldots, J\} \equiv J$. The public signal, denoted y, is drawn from a finite set, Y. The probability that y is realized under the action profile (i, j) is given by ρ_{ij}^{y}. The ex post stage-game payoff to player 1 from the action profile (i, j) and signal y is given by $f_1(i, j, y)$. The ex ante stage-game payoff for player 1 is then $\pi_1(i, j) = \sum_y f_1(i, j, y)\,\rho_{ij}^{y}$. The ex post stage-game payoff to player 2 from the action j and signal y is given by $f_2(j, y)$, and the ex ante stage payoff for player 2 is $\pi_2(i, j) = \sum_y f_2(j, y)\,\rho_{ij}^{y}$.

The stage game is infinitely repeated. Player 1 ("she") is a long-lived (equivalently, long-run) player with discount factor $\delta < 1$. Her payoffs in the infinite horizon game are the average discounted sum of stage-game payoffs, $(1 - \delta) \sum_{t=0}^{\infty} \delta^t \pi_1(i_t, j_t)$. The role of player 2 ("he") is played by a sequence of short-lived (or short-run) players, each of whom only plays once.

The actions of player 2 are public, while player 1's actions are private. Player 1 in period t has a *private history*, consisting of the public signals and all past actions, denoted by $h_{1t} \equiv ((i_0, j_0, y_0), (i_1, j_1, y_1), \ldots, (i_{t-1}, j_{t-1}, y_{t-1})) \in H_{1t} \equiv (I \times J \times Y)^t$. Let $\{\mathcal{H}_{1t}\}_{t=0}^{\infty}$ denote the filtration on $(I \times J \times Y)^{\infty}$ induced by the private histories of player 1. The *public history*, observed by both players, is the sequence $((j_0, y_0), (j_1, y_1), \ldots, (j_{t-1}, y_{t-1})) \in (J \times Y)^t$. Let $\{\mathcal{H}_t\}_{t=0}^{\infty}$ denote the filtration induced by the public histories.

We assume the public signals have full support (Assumption 1), so every signal y is possible after any action profile. We also assume that with sufficient observations player 2 can correctly identify, from the frequencies of the signals, any fixed stage-game action of player 1 (Assumption 2).

Assumption 1 (Full Support): *$\rho_{ij}^y > 0$ for all $(i, j) \in I \times J$ and $y \in Y$.*

Assumption 2 (Identification): *For all $j \in J$, the I columns in the matrix $(\rho_{ij}^y)_{y \in Y, i \in I}$ are linearly independent.*

A behavior strategy for player 1 is a map $\sigma_1 : \cup_{t=0}^{\infty} H_{1t} \to \Delta^I$, from the set of private histories of lengths $t = 0, 1, \ldots$ to the set of distributions over current actions. Similarly, a behavior strategy for player 2 is a map $\sigma_2 : \cup_{t=0}^{\infty} H_t \to \Delta^J$.

A strategy profile $\sigma = (\sigma_1, \sigma_2)$ induces a probability distribution P^{σ} over $(I \times J \times Y)^{\infty}$. Let $E^{\sigma}[\cdot|\mathcal{H}_{\ell t}]$ denote player ℓ's expectations with respect to this distribution conditional on $\mathcal{H}_{\ell t}$, where $\mathcal{H}_{2t} = \mathcal{H}_t$.[7]

In equilibrium, the short-run player plays a best response after every equilibrium history. Player 2's strategy σ_2 is a best response to σ_1 if, for all t,

$$E^{\sigma}[\, \pi_2(i_t, j_t) \mid \mathcal{H}_t] \geq E^{\sigma}[\, \pi_2(i_t, j) \mid \mathcal{H}_t], \qquad \forall j \in J \quad P^{\sigma}\text{-a.s.}$$

Denote the set of such best responses by $BR(\sigma_1)$.

The definition of a Nash equilibrium is completed by the requirement that player 1's strategy maximizes her expected utility:

Definition 1: A *Nash equilibrium of the complete-information game* is a strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$ with $\sigma_2^* \in BR(\sigma_1^*)$ such that for all σ_1:

$$E^{\sigma^*}\left[(1-\delta)\sum_{s=0}^{\infty}\delta^s \pi_1(i_s, j_s)\right] \geq E^{(\sigma_1, \sigma_2^*)}\left[(1-\delta)\sum_{s=0}^{\infty}\delta^s \pi_1(i_s, j_s)\right].$$

The assumption of full-support monitoring ensures that all finite sequences of public signals occur with positive probability, and hence must be

[7] This expectation is well-defined, since I, J, and Y are finite.

followed by optimal behavior in any Nash equilibrium. The only public out-of-equilibrium events are those in which player 2 deviates. Since player 2 is a short-run player, he can never benefit from such a choice. Consequently, any Nash equilibrium outcome is also the outcome of a perfect Bayesian equilibrium.

3.2 The Incomplete-Information Game

At time $t = -1$ a type of player 1 is selected. With probability $1 - p_0 > 0$ she is the "normal" type, denoted by n, with the preferences described above. With probability $p_0 > 0$ she is a "commitment" type, denoted by c, who plays the same (possibly mixed) action $\varsigma_1 \in \Delta^I$ in each period independent of history.[8] We assume:

Assumption 3: *Player 2 has a unique best reply to ς_1 (denoted ς_2) and $\varsigma \equiv (\varsigma_1, \varsigma_2)$ is not a stage-game Nash equilibrium.*

Denote by $\hat{\sigma}_1$ the repeated-game strategy of playing $\varsigma_1 \in \Delta^I$ in each period independent of history. Since ς_2 is the unique best response to ς_1, $BR(\hat{\sigma}_1)$ is the singleton $\{\hat{\sigma}_2\}$, where $\hat{\sigma}_2$ is the strategy of playing ς_2 in each period independent of history. Since ς is not a stage game Nash equilibrium, $(\hat{\sigma}_1, \hat{\sigma}_2)$ is not a Nash equilibrium of the complete-information infinite horizon game.

The example from Section 2 illustrates the role of the assumption that player 2 have a unique best response. The strategy that places equal probability on T and B (while not part of an equilibrium of the stage game) is part of many equilibria of the complete-information game (as long as $\delta > 1/[2(p-q)]$), and consequently the normal type can have a permanent reputation for playing like that commitment type. On the other hand, player 2 has a unique best response to any mixture in which player 1 randomizes with probability of T strictly larger than $\frac{1}{2}$, and a strategy that always plays such a mixture is not part of any equilibrium of the complete-information game.

A state of the world is now a type for player 1 and sequence of actions and signals. The set of states is $\Omega = \{n, c\} \times (I \times J \times Y)^\infty$. The prior p_0, commitment strategy $\hat{\sigma}_1$ and the strategy profile of the normal players $\tilde{\sigma} = (\tilde{\sigma}_1, \sigma_2)$ induce a probability measure P over Ω, which describes how

[8]When we are interested in "Stackelberg" commitment types, and the attendant lower bounds on player 1's ex ante payoffs, it suffices to consider commitment types who follow such simple strategies when player 2 is a short-run type. More complicated commitment types are discussed in Section 6.2.

an uninformed player expects play to evolve. The strategy profile $\hat{\sigma} = (\hat{\sigma}_1, \sigma_2)$ (respectively, $\tilde{\sigma} = (\tilde{\sigma}_1, \sigma_2)$) determines a probability measure $\hat{P}$ (resp., $\tilde{P}$) over Ω, which describes how play evolves when player 1 is the commitment (resp., normal) type. Since $\tilde{P}$ and $\hat{P}$ are absolutely continuous with respect to P, any statement that holds P-almost surely, also holds $\tilde{P}$- and $\hat{P}$-almost surely. Henceforth, we will use $E[\cdot]$ to denote unconditional expectations taken with respect to the measure P. $\tilde{E}[\cdot]$ and $\hat{E}[\cdot]$ are used to denote conditional expectations taken with respect to the measures $\tilde{P}$ and $\hat{P}$. Generic outcomes are denoted by ω. The filtrations $\{\mathcal{H}_{1t}\}_{t=0}^{\infty}$ and $\{\mathcal{H}_t\}_{t=0}^{\infty}$ on $(I \times J \times Y)^{\infty}$ can also be viewed as filtrations on Ω in the obvious way; we use the same notation for these filtrations (the relevant sample space will be obvious).

For any repeated-game behavior strategy $\sigma_1 : \cup_{t=0}^{\infty} H_{1t} \to \Delta^I$, denote by σ_{1t} the t^{th} period behavior strategy, so that σ_1 can be viewed as the sequence of functions $(\sigma_{10}, \sigma_{11}, \sigma_{12}, \ldots)$ with $\sigma_{1t} : H_{1t} \to \Delta^I$. We extend σ_{1t} from H_{1t} to Ω in the obvious way , so that $\sigma_{1t}(\omega) \equiv \sigma_{1t}(h_{1t}(\omega))$, where $h_{1t}(\omega)$ is player 1's t-period history under ω. A similar comment applies to σ_2.

Given the strategy σ_2, the normal type has the same objective function as in the complete-information game. Player 2, on the other hand, is maximizing $E[\ \pi_2(i_t, j) \mid \mathcal{H}_t]$, so that after any history h_t, he is updating his beliefs over the type of player 1 that he is facing. The profile $(\tilde{\sigma}_1, \sigma_2)$ is a Nash equilibrium of the incomplete-information game if each player is playing a best response.

At any equilibrium, player 2's posterior belief in period t that player 1 is the commitment type is given by the $\mathcal{H}_t$-measurable random variable $p_t : \Omega \to [0, 1]$. By Assumption 1, Bayes' rule determines this posterior after all sequences of signals. Thus, in period t, player 2 is maximizing

$$p_t \pi_2(\varsigma_1, j) + (1 - p_t)\, \tilde{E}[\ \pi_2(i_t, j) \mid \mathcal{H}_t]$$

P-almost surely. At any Nash equilibrium of this game, the belief p_t is a bounded martingale with respect to the filtration $\{\mathcal{H}_t\}_t$ and measure P.[9] It therefore converges P-almost surely (and hence $\tilde{P}$- and $\hat{P}$-almost surely) to a random variable p_{∞} defined on Ω. Furthermore, at any equilibrium the posterior p_t is a $\hat{P}$-submartingale and a $\tilde{P}$-supermartingale with respect to the filtration $\{\mathcal{H}_t\}_t$.

A final word on notation: The expression $\tilde{E}\,[\sigma_{1t} | \mathcal{H}_s]$ is the standard conditional expectation, viewed as a $\mathcal{H}_s$-measurable random variable on Ω,

[9]These properties are well-known. Proofs for the model with perfect monitoring (which carry over to imperfect monitoring) can be found in Cripps and Thomas (1995).

while $\tilde{E}[\sigma_1(h_{1t})|h_s]$ is the conditional expected value of $\sigma_1(h_{1t})$ (with h_{1t} viewed as a random history) conditional on the observation of the public history h_s.

4 Impermanent Reputations

4.1 Asymptotic Beliefs

Our main result is:

Theorem 1: *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2, and the commitment action ς_1 satisfies Assumption 3. In any Nash equilibrium of the game with incomplete information, $p_t \to 0$ $\tilde{P}$-almost surely.*

The intuition is straightforward: Suppose there is a Nash equilibrium of the incomplete-information game in which both the normal and the commitment type receive positive probability in the limit (on a positive probability set of histories). On this set of histories, player 2 cannot distinguish between signals generated by the two types (otherwise player 2 could ascertain which type he is facing), and hence must believe that the normal and commitment types are playing the same strategies on average. But then player 2 must play a best response to the commitment type. Since the commitment type's behavior is not a best reply for the normal type (to this player-2 behavior), player 1 must eventually find it optimal to *not* play the commitment-type strategy, contradicting player 2's beliefs. The proof is in Section 5.

As we noted in the Introduction, our argument makes critical use of the assumption of full-support imperfect monitoring. However, if monitoring is perfect and the commitment type plays a mixed strategy, the game effectively has imperfect monitoring (as Fudenberg and Levine (1992) observe). For example, in the perfect monitoring version of the game described in Section 2 (so that player 1's action choice is public), if the commitment type randomizes with probability 3/4 on T, then the realized action choice is a noisy signal of the commitment type. Theorem 1 immediately applies to the perfect monitoring case, as long as the commitment type plays a mixed strategy with full support.

4.2 Asymptotic Equilibrium Play

Given Theorem 1, we should expect continuation play to converge to a Nash equilibrium of the complete-information game. Our next theorem confirms this result.

We use the term *continuation game* for the game with initial period in period t, ignoring the period t histories. We use the notation $t' = 0, 1, 2, \ldots$ for a period of play in a continuation game (which may be the original game) and t for the time elapsed prior to the start of the continuation game. A pure strategy for player 1, s_1, is a sequence of maps $s_{1t'} : H_{1t'} \to I$ for $t' = 0, 1, \ldots$. Thus, $s_{1t'} \in I^{H_{1t'}}$ and $s_1 \in I^{\cup_{t'} H_{1t'}} \equiv S_1$, and similarly $s_2 \in S_2 \equiv J^{\cup_{t'} H_{t'}}$. The spaces S_1 and S_2 are countable products of finite sets. We equip each space S_ℓ, $\ell = 1, 2$, with the σ-algebra generated by the cylinder sets, denoted by $\mathcal{S}_\ell$. The players' payoffs in the infinitely repeated game (as a function of pure strategies) are given by

$$\begin{aligned} u_1(s_1, s_2) &\equiv E[(1-\delta)\sum_{t'=0}^{\infty} \delta^{t'} \pi_1(i_{t'}, j_{t'})], \text{ and} \\ u_2^{t'}(s_1, s_2) &\equiv E[\pi_2(i_{t'}, j_{t'})]. \end{aligned}$$

The expectation above is taken over the action pairs $(i_{t'}, j_{t'})$. These are random, given the pure strategy profile (s_1, s_2), because the pure action played in period t depends upon the random public signals.

For $\ell = 1, 2$, let $\mathcal{M}_\ell$ denote the space of probability measures μ_ℓ on $(S_\ell, \mathcal{S}_\ell)$. We say a sequence of measures $\mu_1^n \in \mathcal{M}_1$ *converges* to $\mu_1 \in \mathcal{M}_1$ if, for each τ, we have

$$\mu_1^n|_{I^{(I\times J\times Y)^\tau}} \to \mu_1|_{I^{(I\times J\times Y)^\tau}} \tag{3}$$

and a sequence of measures $\mu_2^n \in \mathcal{M}_2$ *converges* to $\mu_2 \in \mathcal{M}_2$ if, for each τ, we have

$$\mu_2^n|_{J^{(J\times Y)^\tau}} \to \mu_2|_{J^{(J\times Y)^\tau}}. \tag{4}$$

Moreover, each $\mathcal{M}_\ell$ is sequentially compact in the topology of this convergence. Payoffs for players 1 and 2 are extended to $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ in the obvious way. Since player 1's payoffs are discounted, the inherited product topology is strong enough to guarantee continuity of $u_1 : \mathcal{M} \to \mathbb{R}$. Each player 2's payoff is trivially continuous.

Fix an equilibrium of the incomplete-information game. If the normal type of player 1 observes a private history $h_{1t} \in H_{1t}$, her strategy $\tilde{\sigma}_1$ specifies a behavior strategy in the continuation game. This behavior strategy is realization equivalent to a mixed strategy $\tilde{\mu}_1^{h_{1t}} \in \mathcal{M}_1$ for the continuation game. For a given public history, h_t, there are many possible such mixed strategies that the normal type could be playing. We let $\tilde{\mu}_1^{h_t}$ denote the expected value of $\tilde{\mu}_1^{h_{1t}}$, conditional on the public history h_t. From the point of view of player 2, who observes only the public history, $\tilde{\mu}_1^{h_t}$ is the strategy of the normal player 1 following history h_t. We let $\mu_2^{h_t} \in \mathcal{M}_2$ denote player 2's mixed strategy in the continuation game.

If we had metrized $\mathcal{M}$, a natural formalization of the idea that asymptotically the normal type and player 2 are playing a Nash equilibrium is that the distance between the set of Nash equilibria and the induced distributions $(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})$ goes to zero. While $\mathcal{M}$ is metrizable, a simpler and equivalent formulation is that the limit of every convergent subsequence of $(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})$ is a Nash equilibrium.[10] Section A.1 proves:

THEOREM 2: *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2, and the commitment action ς_1 satisfies Assumption 3. For any Nash equilibrium of the incomplete-information game and for $\tilde{P}$-almost all sequences of histories $\{h_t\}_t$, every cluster point of the sequence of continuation profiles $\{(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})\}_t$ is a Nash equilibrium of the complete-information game with normal player* 1.

Suppose the Stackelberg payoff is not a Nash equilibrium payoff of the complete-information game. Recall that Fudenberg and Levine (1992) provide a lower bound on equilibrium payoffs in the incomplete-information game of the following type: Fix the prior probability of the Stackelberg (commitment) type. Then, there is a value for the discount factor, $\bar{\delta}$, such that if $\delta > \bar{\delta}$, then in every Nash equilibrium, the long-lived player's ex ante payoff is essentially no less than the Stackelberg payoff. The reconciliation of this result with Theorem 2 lies in the order of quantifiers: while Fudenberg and Levine fix the prior, p_0, and then select $\bar{\delta}(p_0)$ large (with $\bar{\delta}(p_0) \to 1$ as $p_0 \to 0$), we fix δ and examine asymptotic play, so that eventually p_t is sufficiently small that $\bar{\delta}(p_t) > \delta$.

4.3 Asymptotic Restrictions on Behavior

This section provides a partial converse to Theorem 2. We identify a class of equilibria of the complete-information game to which (under a continuity hypothesis) equilibrium play of the incomplete-information game can converge.[11] The proof is in Section A.2.

[10] This equivalence is an implication of the sequential compactness of $\mathcal{M}$, since every subsequence of $(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})$ has a convergent sub-subsequence.

[11] We conjecture this hypothesis is redundant, given the other conditions of the theorem, but have not been able to prove it.

Recall that in the example of Section 2, the stage game has a (unique) strict Nash equilibrium BR. It is a straightforward implication of Fudenberg and Levine (1992) that the presence of the commitment type ensures that, as long as player 1 is sufficiently patient, *every* equilibrium in this example begins with a long sequence of play close to TL. On the other hand, an implication of Theorem 3 below is that for the *same* parameters (in particular, the same prior probability of the commitment type), there is an equilibrium in which, with arbitrarily high probability under $\tilde{P}$, BR is eventually played in every period.

The construction of such an equilibrium must address the following two issues. First, as we just observed, reputation effects may ensure that for a long period of time, equilibrium play will be very different from BR. Theorem 3 is consistent with this, since it only claims that in the equilibrium of interest, BR is *eventually* played in every period with high probability. Second, even if reputation effects are not currently operative (because the current belief that player 1 is the commitment type is low), with positive probability (albeit small), a sequence of signals will arise that increases the posterior that player 1 is the commitment type and hence makes reputation effects a recurring possibility.

THEOREM 3: *Suppose the monitoring distribution ρ satisfies Assumptions 1 and 2, and the commitment action ς_1 satisfies Assumption 3. Suppose the stage game has a strict Nash equilibrium, (i^*, j^*). Suppose that for all $\varepsilon > 0$, there exists η and an equilibrium of the complete-information game, $\sigma(0)$, such that for all $p_0 \in (0, \eta)$, the game with incomplete-information with prior p_0 has an equilibrium whose payoff to player 1 is within ε of $u_1(\sigma(0))$. Given any prior p_0 and any δ, for all $\varepsilon > 0$, there exists a Nash equilibrium of the incomplete-information game in which the $\tilde{P}$-probability of the event that eventually (i^*, j^*) is played in every period is at least $1 - \varepsilon$.*

5 PROOF OF THEOREM 1

5.1 Player 2's Posterior Beliefs

The first step is to show that either player 2's expectation (given the public history) of the strategy played by the normal type is in the limit identical to the strategy played by the commitment type, or player 2's posterior probability that player 1 is the commitment type converges to zero (given that player 1 is indeed normal).

This is an extension of a familiar merging-style argument to the case of imperfect monitoring. If the distributions generating player 2's observations are different for the normal and commitment types, then he will be updating his posterior, continuing to do so as the posterior approaches zero. His posterior converges to something strictly positive only if the distributions generating these observations are in the limit identical. In the statement of the following Lemma, h_{1t} is to be interpreted as a function from Ω to $(I \times Y)^t$.

Lemma 1: *At any Nash equilibrium of a game satisfying Assumptions 1 and 2,*[12]

$$\lim_{t\to\infty} p_t(1-p_t)\left\|\varsigma_1 - \tilde{E}[\,\tilde{\sigma}_{1t} \mid \mathcal{H}_t\,]\right\| = 0, \qquad P\text{-a.s.} \tag{5}$$

Proof: Let $p_{t+1}(h_t; j_t, y_t)$ denote player 2's belief in period $t+1$ after playing j_t in period t, observing the signal y_t in period t, and given the history h_t. By Bayes' rule,

$$p_{t+1}(h_t; j_t, y_t) = \frac{p_t \Pr[y_t \mid h_t, j_t, c]}{p_t \Pr[y_t \mid h_t, j_t, c] + (1-p_t)\Pr[y_t \mid h_t, j_t, n]}.$$

The probability player 2 assigns to observing the signal y_t from the commitment type is $\sum_{i\in I} \varsigma_1^i \rho_{ij_t}^{y_t}$, and from the normal type is $\tilde{E}[\sum_{i\in I} \tilde{\sigma}_1^i(h_{1t})\rho_{ij_t}^{y_t} | h_t]$. Using the linearity of the expectations operator, we write $p_{t+1}(h_t; j_t, y_t)$ as

$$p_{t+1}(h_t; j_t, y_t) = \frac{p_t \sum_{i\in I} \rho_{ij_t}^{y_t} \varsigma_1^i}{\sum_{i\in I} \rho_{ij_t}^{y_t}\left(p_t\varsigma_1^i + (1-p_t)\tilde{E}[\tilde{\sigma}_1^i(h_{1t})|h_t]\right)}.$$

Rearranging,

$$\frac{p_{t+1}}{p_t}\sum_{i\in I} \rho_{ij_t}^{y_t}\left(p_t\varsigma_1^i + (1-p_t)\tilde{E}[\tilde{\sigma}_1^i(h_{1t})|h_t]\right) = \sum_{i\in I} \rho_{ij_t}^{y_t}\varsigma_1^i. \tag{6}$$

Denote the summation on the left by A and note that $A < \max_i \rho_{ij_t}^{y_t} < 1$. Repeating the derivation of (6) for $1 - p_{t+1}$, the probability that player 1 is normal, gives $(1-p_{t+1})A/(1-p_t) = \sum_{i\in I} \rho_{ij_t}^{y_t} \tilde{E}[\tilde{\sigma}_1^i(h_{1t})|h_t]$. Taking the difference of this expression and (6) yields

$$A\left|\frac{p_{t+1}}{p_t} - \frac{1-p_{t+1}}{1-p_t}\right| = \left|\sum_{i\in I} \rho_{ij_t}^{y_t}\left(\varsigma_1^i - \tilde{E}[\tilde{\sigma}_1^i(h_{1t})|h_t]\right)\right|.$$

[12] We use $\|x\|$ to denote the sup-norm on $\mathbb{R}^I$.

As $A < 1$ for any fixed realization y of the signal y_t, it follows that for all h_t and j_t,

$$\text{(7)} \quad \max_y |p_{t+1}(h_t; j_t, y) - p_t(h_t)| \geq \\ p_t(h_t)(1 - p_t(h_t)) \left| \sum_{i \in I} \rho^y_{ij_t} \left(\varsigma^i_1 - \tilde{E}[\tilde{\sigma}^i_1(h_{1t})|h_t] \right) \right|.$$

Since p_t is a P-almost sure convergent sequence, it is Cauchy P-almost surely.[13] So the right hand side of (7) converges to zero P-almost surely. Thus, for any y,

$$\text{(8)} \qquad p_t(1 - p_t) \left| \sum_{i \in I} \rho^y_{ij_t} \left(\varsigma^i_1 - \tilde{E}[\tilde{\sigma}^i_{1t}|\mathcal{H}_t] \right) \right| \to 0, \qquad P\text{-a.s.}$$

Hence, if both types are given positive probability in the limit then the frequency that any signal is observed is identical under the two types.

We now show that (8) implies (5). Let Π_{j_t} be a $|Y| \times |I|$ matrix whose y^{th} row, for each signal $y \in Y$, contains the terms $\rho^y_{ij_t}$ for $i = 1, \ldots, |I|$. Then as (8) holds for all y (and Y is finite), it can be restated as

$$\text{(9)} \qquad p_t(1 - p_t) \left\| \Pi_{j_t} \left(\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_t] \right) \right\| \to 0, \qquad P\text{-a.s.},$$

where $\|.\|$ is the supremum norm. By Assumption 2, the matrices Π_{j_t} have I linearly independent columns for all j_t, so $x = 0$ is the unique solution to $\Pi_{j_t} x = 0$ in $\mathbb{R}^I$. In addition, there exists a strictly positive constant $b = \inf_{j \in J, x \neq 0} \|\Pi_j x\| / \|x\|$. Hence $\|\Pi_j x\| \geq b\|x\|$ for all $x \in \mathbb{R}^I$ and all $j \in J$. From (9), we then get

$$p_t(1 - p_t) \left\| \Pi_{j_t} \left(\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_t] \right) \right\| \\ \geq p_t(1 - p_t) b \left\| \varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_t] \right\| \to 0, \qquad P\text{-a.s.},$$

which implies (5). *Q.E.D.*

Condition (5) says that either player 2's best prediction of the normal type's behavior is eventually identical to the commitment type's behavior (that is, $\| \varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t} \mid \mathcal{H}_t] \| \to 0$ P-almost surely), or the type is revealed

[13] Note that the analysis is now *global*, rather than *local*, in that we treat all the expressions as functions on Ω.

(that is, $p_\infty(1-p_\infty) = 0$ P-almost surely). However, $p_\infty < 1$ $\tilde{P}$-almost surely, and hence (5) implies a simple corollary:[14]

Corollary 1: *At any equilibrium of a game satisfying Assumptions 1 and 2,*

$$\lim_{t\to\infty} p_t \left\| \varsigma_1 - \tilde{E}[\ \tilde{\sigma}_{1t} \mid \mathcal{H}_t\ ] \right\| = 0, \qquad \tilde{P}\text{-a.s.}$$

5.2 Beliefs about Player 2's Beliefs

Corollary 1 implies that if $p_t \not\to 0$ on a set of states with positive measure, then (on this set of states) player 2 must think that the normal type's strategy is arbitrarily close to that of the commitment type. Since player 1 is better informed than player 2, player 1 must know that player 2 believes this:

Lemma 2: *Suppose Assumptions 1 and 2 are satisfied. Suppose there exists $A \subset \Omega$ such that $\tilde{P}(A) > 0$ and $p_\infty(\omega) > 0$ for all $\omega \in A$. Then, for sufficiently small $\eta > 0$, there exists $F \subset A$ with $\tilde{P}(F) > 0$ such that, for any $\xi > 0$, there exists T for which*

$$p_t > \eta, \qquad \forall t > T \tag{10}$$

and

$$\tilde{E}\left[\sup_{s\geq t} \left\| \varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s] \right\| \ \middle|\ \mathcal{H}_t \right] < \xi, \qquad \forall t > T \tag{11}$$

for all $\omega \in F$ and, for all $\psi > 0$,

$$\tilde{P}\left(\sup_{s\geq t} \left\| \varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s] \right\| < \psi \ \middle|\ \mathcal{H}_t \right) \to 1, \tag{12}$$

where the convergence is uniform on F.

Proof: Define the event $D_\eta = \{\omega \in A : p_\infty(\omega) > 2\eta\}$. Because the set A on which $p_\infty(\omega) > 0$ has $\tilde{P}$-positive measure, for any $\eta > 0$ sufficiently small, we have $\tilde{P}(D_\eta) > 2\mu$, for some $\mu > 0$. On the set of states D_η the random variable $\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_t]\|$ tends $\tilde{P}$-almost surely to zero (by Lemma

[14] Since the odds ratio $p_t/(1-p_t)$ is a $\tilde{P}$-martingale, $p_0/(1-p_0) = \tilde{E}[p_t/(1-p_t)]$ for all t. The left side of this equality is finite, so $\lim p_t < 1$ $\tilde{P}$-almost surely.

1). Therefore, on D_η the random variable $Z_t = \sup_{s\geq t} \|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\|$ converges $\tilde{P}$-almost surely to zero and hence (Hart (1985, Lemma 4.24))

$$\tilde{E}[\ Z_t \mid \mathcal{H}_t\ ] \to 0 \quad \tilde{P} - \text{almost surely.} \tag{13}$$

Egorov's Theorem (Chung (1974, p.74)) then implies that there exists $F \subset D_\eta$ such that $\tilde{P}(F) \geq \mu$ on which the convergence of p_t and $\tilde{E}[Z_t|\mathcal{H}_t]$ is uniform. Hence, for any $\xi > 0$, there exists a time T such that the inequalities in (10) and (11) hold everywhere on F for all $t > T$.

Fix $\psi > 0$. Then, for all $\xi' > 0$, (11) holds for $\xi = \xi'\psi$, which implies that, uniformly on F,

$$\tilde{P}\left(\sup_{s\geq t}\left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \psi \;\middle|\; \mathcal{H}_t\right) \to 1.$$

Q.E.D.

5.3 Completion of the Proof of Theorem 1

The proof of Theorem 1 is completed by showing that, on a subset of the states F in Lemma 2, player 2 believes he should be playing a best response to the commitment strategy. The normal type of player 1 will best respond to this player-2 best response with high probability, ensuring that the normal and commitment types of player 1 play differently, contradicting the assumption that $p_t \not\to 0$ on F.

Define $\beta \equiv \min_i\{\varsigma_1^i : \varsigma_1^i > 0\}$ and $\gamma \equiv \min_{y,i,j} \rho_{ij}^y$, where the latter is strictly positive by Assumption 1. Since ς_1 is not a best reply for the normal type to ς_2 (the myopic best reply to ς_1), there exists $\eta > 0$ such that for any repeated-game strategy for player 2 that attaches probability at least $1-\eta$ to $\hat{\sigma}_2$ (i.e., to always playing ς_2), ς_1 is suboptimal for the normal type in period 1.

As ς_2 is the unique best response to ς_1, it is strict and so there exists $\psi > 0$ such that ς_2 is the unique best response to any action of player 1, ς_1', satisfying $\|\varsigma_1' - \varsigma_1\| < \psi$.

Suppose that there is a positive $\tilde{P}$-probability set of outcomes A on which $p_\infty > 0$. Choose ξ, ζ such that $\zeta < \beta\gamma$ and $\xi < \min\{\psi, \beta - \zeta\gamma\}$. By (12), there is a $\tilde{P}$-positive measure set $F \subset A$ and T such that, on F and for any $t > T$,

$$\tilde{P}\left(\sup_{s\geq t}\left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \xi \;\middle|\; \mathcal{H}_t\right) > 1 - \eta\zeta. \tag{14}$$

Hence, on F,

$$\left\| \varsigma_1 - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_t] \right\| < \xi \qquad \tilde{P} \text{ a.s.} \tag{15}$$

Set

$$g_t \equiv \tilde{P}\left(\sup_{s \geq t} \left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \xi \;\middle|\; \mathcal{H}_{1t}\right)$$

and $\kappa_t \equiv \tilde{P}(g_t > 1 - \eta \mid \mathcal{H}_t)$. As $\{\mathcal{H}_{1t}\}_t$ is a finer filtration than $\{\mathcal{H}_t\}_t$,

$$\begin{aligned} \tilde{P}\left(\sup_{s \geq t} \left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \xi \;\middle|\; \mathcal{H}_t\right) &= \tilde{E}[\ g_t \mid \mathcal{H}_t\ ] \\ &= \tilde{E}\left[g_t \mid g_t \leq 1 - \eta, \mathcal{H}_t\right](1 - \kappa_t) + \tilde{E}\left[g_t \mid g_t > 1 - \eta, \mathcal{H}_t\right]\kappa_t \\ &\leq (1 - \eta)(1 - \kappa_t) + \kappa_t. \end{aligned} \tag{16}$$

Combining the inequalities (14) and (16) we get that for almost every state in F, $\kappa_t > 1 - \zeta$. That is, for all $t > T$ and for almost every state in F,

$$\tilde{P}\left(\tilde{P}\left(\sup_{s \geq t} \left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \xi \middle| \mathcal{H}_{1t}\right) > 1 - \eta \;\middle|\; \mathcal{H}_t\right) > 1 - \zeta, \tag{17}$$

and so player 2 assigns a probability of at least $1 - \zeta$ to player 1 believing with probability at least $1 - \eta$ that player 2 believes player 1's strategy is within ξ of the commitment strategy.

Since $\xi < \psi$, player 2 plays ς_2, the unique best response to the commitment action, whenever he believes that 1's strategy is within ξ of the commitment strategy. Hence, in any period $t > T$, player 2 assigns a probability of at least $1 - \zeta$ to player 1 believing that player 2's subsequent play is $\hat{\sigma}_2$ with at least probability $1 - \eta$. Thus, player 2 assigns probability at least $1 - \zeta$ to player 1's subsequent play being a best response to player 2's best response to $\hat{\sigma}_1$. But η was chosen so that there is then an action in the support of $\hat{\sigma}_1$, say i', that is not optimal in period t. Player 2 must accordingly believe that i' is played with a probability of no more than ζ in period t. But since $\beta - \zeta > \xi$, this contradicts (15). *Q.E.D.*

6 Extensions

6.1 Many Commitment Types

To extend the preceding analysis to the case in which there are many commitment types, let $\mathcal{T}$ be a set of possible commitment types. The commitment type c plays the repeated-game strategy $\hat{\sigma}_1^c$ that plays the fixed stage-game action $\varsigma_1^c \in \Delta^I$ in each period. We assume $\mathcal{T}$ is either finite or countably infinite, and $\varsigma_1^c \neq \varsigma_1^{c'}$ for all $c \neq c' \in \mathcal{T}$. At time $t = -1$ a type of player 1 is selected. With probability $p_0^c > 0$, she is commitment type c, and with probability $p_0^n = 1 - \sum_{c \in \mathcal{T}} p_0^c > 0$ she is the "normal" type. A state of the world is, as before, a type for player 1 and sequence of actions and signals. The set of states is then $\Omega = \mathcal{T} \times (I \times J \times Y)^\infty$. We denote by $\hat{P}^c$ the probability measure induced on Ω by the commitment type $c \in \mathcal{T}$, and as usual, we denote by $\tilde{P}$ the probability measure on Ω induced by the normal type. Finally, we denote by p_t^c player 2's period t belief that player 1 is the commitment type c.

To deal with many types of player 1, we first argue that it is impossible for two different commitment types to be given positive probability in the limit.

LEMMA 3: *At any Nash equilibrium of a game satisfying Assumptions 1 and 2, for all $c \neq c' \in \mathcal{T}$,*

$$p_t^c p_t^{c'} \to 0 \qquad P\text{−a.s.}$$

PROOF: Derive (6) for each of the types c and c'. Take the difference of these two equations, repeat the remaining part of the proof of Lemma 1, and use $\varsigma_1^c \neq \varsigma_1^{c'}$. *Q.E.D.*

THEOREM 4: *Suppose ρ satisfies Assumptions 1 and 2. Let $\mathcal{T}^*$ be the set of commitment types $c \in \mathcal{T}$ for which player 2 has a unique best response ς_2^c to ς_1^c, with $(\varsigma_1^c, \varsigma_2^c)$ not a Nash equilibrium of the stage game. Then in any Nash equilibrium, $p_t^c \to 0$ for all $c \in \mathcal{T}^*$ $\tilde{P}$-almost surely.*

The proof duplicates that of Theorem 1, with the following change. Fix some type $c' \in \mathcal{T}^*$. In the proof, reinterpret $\tilde{P}$ as $P^{-c'} = \sum_{c \neq c'} p_0^c \hat{P}^c + p_0^n \tilde{P}$, the unconditional measure on Ω implied by the normal type and all the commitment types other than c'. The only point at which it is important that $\tilde{P}$ is indeed the measure induced by the normal type is at the end of the proof, when the normal type has a profitable deviation that contradicts player 2's beliefs. We now apply Lemma 3. Since we are arguing on a $\tilde{P}$-positive probability subset where $p_t^{c'}$ is not converging to zero, every other commitment type is receiving little weight in 2's beliefs. Consequently, from player 2's point of view, eventually the measures $P^{-c'}$ and $\tilde{P}$ are sufficiently close to obtain the same contradiction.

6.2 Complicated Commitment Types

We have followed the common practice of considering *simple* commitment types who repeat a fixed stage-game mixture in each period. The results extend to commitment types whose strategies are not stationary, as long as their behavior is *eventually* incompatible with equilibrium.

Definition 2: The strategy $\bar{\sigma}_1$ is *never an equilibrium strategy in the long run,* if there exists $\bar{T}$ and $\varepsilon > 0$ such that, for every $\bar{\sigma}_2 \in BR(\bar{\sigma}_1)$ and for every $t \geq \bar{T}$, there exists $\tilde{\sigma}_1$ such that $P^{\bar{\sigma}}$-a.s,

$$E^{\bar{\sigma}}\left[(1-\delta)\sum_{s=t}^{\infty}\delta^{s-t}\pi_1(i_s,j_s)\middle|\mathcal{H}_{1t}\right]+\varepsilon \;<\; E^{(\tilde{\sigma}_1,\bar{\sigma}_2)}\left[(1-\delta)\sum_{s=t}^{\infty}\delta^{s-t}\pi_1(i_s,j_s)\middle|\mathcal{H}_{1t}\right].$$

A strategy $\overline{\sigma}_1$ is *simple* if it plays the same stage-game (possibly mixed) action after every history. A strategy $\overline{\sigma}_1$ is *public* if it is measurable with respect to $\{\mathcal{H}_t\}_t$, so that the mixture over actions in each period depends only upon the public history. A strategy $\overline{\sigma}_1$ is *publicly implementable by a finite automaton* if there exists a finite set W, an action function $d : W \to \Delta^I$, a transition function $\varphi : W \times Y \to W$, and an initial element $w_0 \in W$, such that $\sigma_1(h_t) = d(w(h_t))$, where $w(h_t)$ is the state reached from w_0 under the public history h_t and transition rule φ.

It is straightforward to show that if a simple strategy plays the stage-game mixture $\varsigma \in \Delta^I$, to which player 2 has a unique best response, then the strategy is never an equilibrium strategy in the long run if and only if ς is not part of a stage-game Nash equilibrium. Similarly, suppose $\bar{\sigma}_1$ is publicly implementable by the finite automaton (W, d, φ, w_0), with every state in W reachable from every other state in W under φ. If player 2 has a unique best reply to $d(w)$ for all $w \in W$, then $\bar{\sigma}_1$ is never an equilibrium strategy in the long run if and only if $\bar{\sigma}_1$ is not part of a Nash equilibrium of the complete-information game.[15]

Theorem 5: *Suppose ρ satisfies Assumptions 1 and 2. Suppose $\hat{\sigma}_1$ is a public strategy with finite range (i.e., $\cup_{h_t}\hat{\sigma}_1(h_t)$ is finite) that is never an*

[15] The only if direction of this statement is obvious. So, suppose $\bar{\sigma}_1$ is not a Nash equilibrium of the complete-information game. Since player 2 always has a unique best reply to $d(w)$, σ_2 is public, and can also be represented as a finite-state automaton, with the same set of states and transition function as $\bar{\sigma}_1$. Since $\bar{\sigma}_1$ is not a Nash equilibrium, there is some state $w' \in W$, and some action i' not in the support of $d(w')$ such that when the state is w', playing i' and then following $\bar{\sigma}_1$ yields a payoff that is strictly higher than following $\bar{\sigma}_1$ at w'. Since the probability of reaching w' from any other state is strictly positive (and so bounded away from zero), $\bar{\sigma}_1$ is never an equilibrium in the long run.

equilibrium strategy in the long run. In any Nash equilibrium of any game with incomplete information, $p_t \to 0$ $\tilde{P}$-almost surely.

PROOF: Since $\hat{\sigma}_1$ is never an equilibrium strategy in the long run, there exists $\bar{T}$ such that after any positive probability history of length at least $\bar{T}$, $\hat{\sigma}_1$ is not a best response to any strategy $\sigma_2 \in BR(\hat{\sigma}_1)$ of player 2 that best responds to $\hat{\sigma}_1$. Indeed, there exists $\eta > 0$ such that this remains true for any strategy of player 2 that attaches probability at least $1-\eta$ to any strategy in $BR(\hat{\sigma}_1)$.

The argument in Section 5.3 now applies, with the following three changes: First, redefine β as $\beta \equiv \min_{i,h_t}\{\hat{\sigma}_1^i(h_t) : \hat{\sigma}_1^i(h_t) > 0\}$ (which is strictly positive, since $\hat{\sigma}_1$ has finite range). Second, T must be larger than $\bar{T}$. Third, the last two paragraphs of that section are replaced by the following:

We now argue that there is a period $t \geq T$ and an outcome in F such that $\hat{\sigma}_1$ is not optimal for the normal player 1 in period t. Given any outcome $\omega \in F$ and a period $t \geq T$, let h_t be its t-period public history. There is a $K > 0$ such that for any t large, there is a public history $y_t, \ldots, y_{t+k}$, $0 \leq k \leq K$, under which $\hat{\sigma}_1(h_t, y_t, \ldots, y_{t+k})$ puts positive probability on a suboptimal action. (Otherwise, no deviation can increase the period-t expected continuation payoff by at least ε.) Moreover, by full support, any K sequence of signals has probability at least $\lambda > 0$. If the public history $(h_t, y_t, \ldots, y_{t+k})$ is consistent with an outcome in F, then we are done. So, suppose there is no such outcome. That is, for every $t \geq T$, there is no outcome in F for which $\hat{\sigma}_1$ attaches positive probability to a suboptimal action within the next K periods. Letting $C_t(F)$ denote the t-period cylinder set of F, $\tilde{P}(F) \leq \tilde{P}(C_{t+K}(F)) \leq (1-\lambda)\,\tilde{P}(C_t(F))$ (since the public history of signals that leads to a suboptimal action has probability at least λ). Proceeding recursively from T, we have $\tilde{P}(F) \leq \tilde{P}(C_{T+\ell K}(F)) \leq (1-\lambda)^\ell\,\tilde{P}(C_T(F))$, and letting $\ell \to \infty$, we have $\tilde{P}(F) = 0$, a contradiction.

Hence, there is a period $t \geq T$ and an outcome in F such that one of the actions in the support of $\hat{\sigma}_1$, i' say, is not optimal in period t. That is, any best response assigns zero probability to i' in period t. From (17), player 2's beliefs give a probability of at least $1-\zeta$ to a strategy of player 1 that best responds to 2's best response to $\hat{\sigma}_1$, which means that player 2 believes that i' is played with a probability of no more than ζ. But since $\beta - \zeta > \xi$, this contradicts (15). *Q.E.D.*

6.3 Two Long-Lived Players

We now extend the analysis to the case of a long-lived player 2. The second and third paragraphs of Section 5.3 are the only places where the assumption that player 2 is short-lived makes an appearance. When player 2 is short-lived, player 2 is myopically best responding to the current play of player 1, and so as long as player 2 is sufficiently confident that he is facing the commitment type, he will best respond to the commitment type. On the other hand, if player 2 is long-lived, like player 1, then there is no guarantee that this is still true. For example, player 2 may find experimentation profitable. Nonetheless, reputation effects can still be present (Celentani, Fudenberg, Levine, and Pesendorfer (1996)).

The following result (proven in the Appendix) shows that if the commitment type and the normal type are behaving sufficiently similarly, then player 2 will be playing a best response to the commitment type for arbitrarily many periods. (The notation (W, d, φ, w_0) is described above Theorem 5.)

LEMMA 4: *Suppose $\hat{\sigma}_1$ is publicly implementable by the finite automaton (W, d, φ, w_0), and $BR(\hat{\sigma}_1; w')$ is the set of best replies for player 2 to (W, d, φ, w'). For any history h_t, let $w(h_t) \in W$ be the state reached from w_0 under the public history consistent with h_t. Let $(\tilde{\sigma}_1, \sigma_2)$ be Nash equilibrium strategies in the incomplete-information game where player 2 is long-lived with discount factor $\delta_2 \in [0, 1)$. Suppose $\exists \kappa > 0$ such that for all h_t, if $\sigma_2^j(h_t) > 0$ then $\sigma_2^j(h_t) > \kappa$. Then, then for all $T > 0$ there exists $\psi > 0$ such that if player 2 observes a history h_t so that*

$$P\left(\sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s}|\mathcal{H}_s] - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s] \right\| < \psi \,\middle|\, h_t\right) > 1 - \psi, \tag{18}$$

then for some $\sigma_2' \in BR(\hat{\sigma}_1; w(h_t))$, the continuation strategy of σ_2 after the history h_t agrees with σ_2' for the next T periods.

If player 2's posterior that player 1 is the commitment type fails to converge to zero on a set of states of positive $\tilde{P}$-measure, then the same argument as in Lemma 2 shows that (18) holds (note that (11) in Lemma 2 uses $\tilde{P}$ rather than P to evaluate the probability of the event of interest).

With this result in hand, the proof of Theorem 1 goes through as before, establishing:

THEOREM 6: *Suppose ρ satisfies Assumptions 1 and 2. Suppose $\hat{\sigma}_1$ is publicly implementable by a finite automaton and is never an equilibrium strategy in the long run. Let $(\tilde{\sigma}_1, \sigma_2)$ be Nash equilibrium strategies in the incomplete-information game where player 2 is long-lived with discount factor $\delta_2 \in [0, 1)$. Suppose $\exists \kappa > 0$ such that for all h_t, if $\sigma_2^j(h_t) > 0$ then $\sigma_2^j(h_t) > \kappa$. Then, $p_t \to 0$ $\tilde{P}$-almost surely.*

6.4 Private Actions

Our results continue to hold when player 2's actions are private, as long as player 1 can infer player 2's posterior belief p_t from the public signals.[16] This will be the case if

$$\rho^y_{ij}\rho^y_{i'j'} = \rho^y_{i'j}\rho^y_{ij'}$$

for all $y \in Y$, $i, i' \in I$, and $j, j' \in J$. This holds, for example, if the public signal y is a vector $(y_1, y_2) \in Y_1 \times Y_2 = Y$, with y_1 a signal of player 1's action and y_2 an independent signal of player 2's action. In this case, action i induces a probability distribution ρ_i over Y_1 while action j induces ρ_j over Y_2, with

$$\rho^y_{ij} = \rho^{y_1}_i \rho^{y_2}_j \qquad \forall i, j, y. \tag{19}$$

The full-support Assumption 1 is replaced by the requirement that, for all i and $y_1 \in Y_1$,

$$\rho^{y_1}_i > 0.$$

Assumption 2, in the presence of (19), is equivalent to the requirement that there are I linearly independent columns in the matrix

$$(\rho^{y_1}_i)_{y_1 \in Y_1, i \in I}\,.$$

Cripps, Mailath, and Samuelson (2003) addresses the case where player 2's actions are not known to player 1 and his posterior depends upon his actions as well as the public signals. In this case, the long-lived player's reputation is *private*, since the public signals do not allow player 1 to infer 2's posterior beliefs. This complicates the analysis, since it is now harder to show that the convergence of player 2's beliefs implies that the normal player 1 knows she has a profitable deviation from the commitment strategy. In the course of coping with the potential uninformativeness of the public signals, we extend the results to the case of purely private monitoring.

Olin School of Business, Washington University in St. Louis, St. Louis, MO 63130-4899; cripps@olin.wustl.edu;

Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297; gmailath@econ.upenn.edu;

and

Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706-1320; LarrySam@ssc.wisc.edu.

[16]Indeed, each player 2's action choices can be completely private, so that future player 2's do not learn the choice of the active player 2.

A Appendix

A.1 Proof of Theorem 2

Proof: At the given equilibrium, the normal type is playing in an optimal way from time t onwards given her (private) information. Thus, for each history h_{1t}, derived public history h_t, and strategy $s_1' \in S_1$,

$$E_{(\tilde{\mu}_1^{h_{1t}}, \mu_2^{h_t})}[u_1(s_1, s_2)] \geq E_{\mu_2^{h_t}}[u_1(s_1', s_2)].$$

The subscripts on the expectation operator are the measures on (s_1, s_2). Moreover, for the derived public history h_t and any strategy $s_1' \in S_1$,

$$E_{(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})}[u_1(s_1, s_2)] \geq E_{\mu_2^{h_t}}[u_1(s_1', s_2)]. \tag{A.1}$$

Player 2 is also playing optimally from time t onwards given the public information, which implies that for all $s_2' \in S_2$, all $h_{t'}$ and all $t' > 0$,

$$E_{(p_t\hat{\mu}_1^{h_t}+(1-p_t)\tilde{\mu}_1^{h_t}, \mu_2^{h_t})}[u_2^{t'}(s_1, s_2)] \geq E_{p_t\hat{\mu}_1^{h_t}+(1-p_t)\tilde{\mu}_1^{h_t}}[u_2^{t'}(s_1, s_2')], \tag{A.2}$$

where $\hat{\mu}_1^{h_t}$ is the play of the commitment type. Since player 2 is a short-run player, this inequality is undiscounted and holds for all t'.

From Theorem 1, $p_t \to 0$ $\tilde{P}$-almost surely. Suppose $\{h_t\}_t$ is a sequence of public histories with $p_t \to 0$, and suppose $\{(\tilde{\mu}_1^{h_t}, \mu_2^{h_t})\}_{t=1}^{\infty} \to (\tilde{\mu}_1^*, \mu_2^*)$ on this sequence. We need to show that $(\tilde{\mu}_1^*, \mu_2^*)$ satisfies (A.1) and (A.2) (the latter for all $t' > 0$). It suffices that the expectations $E_{(\mu_1,\mu_2)}[u_1(s_1, s_2)]$ and $E_{(\mu_1,\mu_2)}[u_2(s_1, s_2)]$ are continuous in (μ_1, μ_2). The continuity required is established in the proof of Theorem 4.4 in Fudenberg and Tirole (1991). *Q.E.D.*

A.2 Proof of Theorem 3

We begin by focusing on games that are "close" to the complete-information game. All the Lemmas assume the hypotheses of Theorem 3.

Lemma A: *For all T, there exists $\hat{\eta} > 0$ such that for all $p_0 \in (0, \hat{\eta})$, there is a Nash equilibrium of the incomplete-information game in which the normal type plays i^* and player 2 plays j^* for the first T periods, irrespective of history.*

Proof: Let $\varepsilon' = \frac{1}{2}\left[\pi_1(i^*, j^*) - \max_{i \neq i^*} \pi_1(i, j^*)\right] > 0$. By assumption, there exists $\eta > 0$ and a Nash equilibrium of the complete-information game,

$\sigma(0)$, such that for each belief $p \in (0, \eta)$, there is a Nash equilibrium of the incomplete-information game, $\sigma(p)$, satisfying $|E_p u_1(\sigma(p)) - E_0 u_1(\sigma(0))| < \frac{\varepsilon'}{2}$, where E_p denotes taking expectations with probability p on the commitment type. Hence, for $p, p' < \eta$, $\left|E_p u_1(\sigma(p)) - E_{p'} u_1(\sigma(p'))\right| < \varepsilon'$.

Since j^* is player 2's strict best response to i^*, there exists $\eta' > 0$ so that for all $p_t < \eta'$, j^* is still a best response to the normal type playing i^*. For any T there exists $\hat{\eta} > 0$ so that if $p_0 < \hat{\eta}$, then $p_t < \min\{\eta, \eta'\}$ in all periods $t \leq T$, by Assumption 1. The equilibrium strategy profile is to play (i^*, j^*) for the first T periods (ignoring history), and then play according to the strategy profile identified in the previous paragraph for the belief p_T, $\sigma(p_T)$. By construction, no player has an incentive to deviate and so the profile is indeed a Nash equilibrium. *Q.E.D.*

While, for T large, the equilibrium just constructed yields payoffs to player 1 that are close to $\pi_1(i^*, j^*)$, the equilibrium guarantees nothing about asymptotic play. The equilibrium of the next Lemma does.

LEMMA B: *For all $\varepsilon > 0$, there exists $\eta^* > 0$ such that for all $p_0 \in (0, \eta^*]$, there is a Nash equilibrium of the incomplete-information game, $\sigma^{**}(p_0)$, in which the $\tilde{P}$-probability of the event that (i^*, j^*) is played in every period is at least $1 - \varepsilon$.*

PROOF: Fix $\zeta = \frac{1}{3}[\pi_1(i^*, j^*) - \max_{i \neq i^*} \pi_1(i, j^*)] > 0$, and choose T large enough so that $\delta^T M < \frac{\zeta}{2}$ (recall that M is an upper bound for stage game payoffs) and that the average discounted payoff to player 1 from T periods of (i^*, j^*) is within $\frac{\zeta}{2}$ of $\pi_1(i^*, j^*)$. Denote by $\hat{\eta}$ the upper bound on beliefs given in Lemma A. For any prior $p \in (0, \hat{\eta})$ that player 1 is the commitment type, let $\sigma^*(p)$ denote the equilibrium of Lemma A. By construction, $\sigma^*(p)$ yields player 1 an expected payoff within ζ of $\pi_1(i^*, j^*)$.

There exists $\eta'' < \hat{\eta}$ such that if $p_t < \eta''$, then the posterior after T periods, $p_{t+T}(p_t)$, is necessarily below $\hat{\eta}$. Consider the following strategy profile, consisting of two phases. In the first phase, play (i^*, j^*) for T periods, ignoring history. In the second phase, behavior depends on the posterior beliefs of player 2, $p_{t+T}(p_t)$. If $p_{t+T}(p_t) > \eta''$, play $\sigma^*(p_{t+T}(p_t))$. If $p_{t+T}(p_t) \leq \eta''$, begin the first phase again.

By construction, the continuation payoffs at the end of the first phase are all within ζ of $\pi_1(i^*, j^*)$, and so for any prior satisfying $p_0 < \eta''$, the strategy profile is an equilibrium.

Fix p_0, and let $p_t^\dagger$ be the beliefs of player 2 under the strategy profile in which (i^*, j^*) is played in every period, irrespective of history. It is immediate that $p_t^\dagger \to 0$ $P^\dagger$-almost surely (where $P^\dagger$ is the measure implied by (i^*, j^*) in every period), and so $\sup_{t' \geq t} p_{t'}^\dagger \to 0$ $P^\dagger$-almost surely. Moreover, if $p_\tau^\dagger \leq \eta''$ for all $\tau \leq t$, then $p_\tau^\dagger = p_\tau$ for all $\tau \leq t$. By Egorov's Theorem, there exists a t^* such that $P^\dagger\{\sup_{t' \geq t^*} p_{t'}^\dagger \leq \eta''\} > 1 - \varepsilon$. But then for some public history, h_{t^*}, $P^\dagger\{\sup_{t' \geq t^*} p_{t'}^\dagger \leq \eta'' | h_{t^*}\} > 1 - \varepsilon$. The monotonicity of $p_t^\dagger$ as a function of p_0 implies that, for some $\eta^* > 0$, if $p_0 < \eta^*$, $p_t^\dagger \leq \eta''$ for all $t \leq t^*$. Moreover, the set $\{\sup_{t' \geq t^*} p_{t'}^\dagger \leq \eta''\}$ cannot shrink as p_0 is reduced, and so $P^\dagger\{\sup_t p_t^\dagger \leq \eta''\} > 1 - \varepsilon$. Hence, for $p_0 < \eta^*$, $\tilde{P}\{\sup_t p_t \leq \eta''\} = P^\dagger\{\sup_t p_t^\dagger \leq \eta''\} > 1 - \varepsilon$. *Q.E.D.*

PROOF OF THEOREM 3:
We first construct an equilibrium of an artificial game, and then argue that this equilibrium induces an equilibrium with the desired properties in the original game.

Fix ε and the corresponding η^* from Lemma B. In the artificial game, player 2 has the action space $J \times \{g, e\} \times [0, 1]$, where we interpret g as "go," e as "end," and $p \in [0, 1]$ as an announcement of the posterior belief of player 2. The game is over immediately when player 2 chooses e. The payoffs for player 2 when player 2 ends the game with the announcement of p depend on the actions as well as on the type of player 1 (recall that n is the normal type and c is the commitment type):

$$\pi_2^*(i, j, e, p; n) = \pi_2(i, j) + \eta^* - p^2$$

and

$$\pi_2^*(i, j, e, p; c) = \pi_2(i, j) - (1 - \eta^*) - (1 - p)^2,$$

where $\eta^* > 0$ is from Lemma B. The payoffs for player 2 while the game continues are:

$$\pi_2^*(i, j, g, p; n) = \pi_2(i, j) - p^2$$

and

$$\pi_2^*(i, j, e, p; c) = \pi_2(i, j) - (1 - p)^2.$$

The payoffs for the normal type of player 1 from the outcome $\{(i_s, j_s, g, p_s)\}_{s=0}^\infty$ (note that player 2 has always chosen g) are as before (in particular, the belief announcements are irrelevant):

$$(1 - \delta) \sum_{s=0}^{\infty} \delta^s \pi_1(i_s, j_s).$$

For the outcome $\left\{(i_s, j_s, g)_{s=0}^{t-1}, (i_t, j_t, e, p_t)\right\}$, the payoffs for player 1 are

$$(1-\delta)\sum_{s=0}^{t}\delta^s \pi_1(i_s, j_s) + \delta^t u_1(\sigma^{**}(p_t)),$$

where $u_1(\sigma^{**}(p_t))$ is player 1's equilibrium payoff under $\sigma^{**}(p_t)$ from Lemma B.

Since player 2 chooses an announcement $p \in [0,1]$ to minimize $(1-p_t)p^2 + p_t(1-p)^2$, he always finds it strictly optimal to announce his posterior. Moreover, again by construction, player 2 ends the game if and only if his posterior is less than η^*. Moreover, the artificial game has an equilibrium (σ_1^*, σ_2^*) (by Fudenberg and Levine (1983, Theorem 6.1)).

The desired equilibrium in the original game is given by (σ_1^*, σ_2^*), with the modification that should (σ_1^*, σ_2^*) call for player 2 to announce e, then play proceeds according to the equilibrium specified in Lemma B for the corresponding value of ρ ($< \eta^*$). It follows from Lemma B that this is an equilibrium of the original game. It then follows from Theorem 1 that $\tilde{P}$-almost surely, the probability of the event that (i^*, j^*) is played eventually is at least $1-\varepsilon$. *Q.E.D.*

A.3 Proof of Lemma 4

Proof: Fix $T > 0$. Since W is finite, it is enough to argue that for each $w \in W$, there is $\psi_w > 0$ such that if player 2 observes a history h_t so that $w = w(h_t)$ and

$$P\left\{\sup_{s\geq t}\left\|\hat{E}[\hat{\sigma}_{1s}|\mathcal{H}_s] - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_s]\right\| < \psi_w \middle| h_t\right\} > 1-\psi_w, \tag{A.3}$$

then for some $\sigma_2' \in BR(\hat{\sigma}_1; w)$, the continuation strategy of σ_2 after the history h_t agrees with σ_2' for the next T periods.

Fix a public history, h_t'. Let $\hat{\sigma}_1(h_s)$ denote the play of the finite automaton $(W, d, \varphi, w(h_t'))$ after the public history h_s, where h_t' is the initial segment of h_s. Since player 2 is discounting, there exists T' such for any $w \in W$, there is $\varepsilon_w > 0$ such that if for $s = t, \ldots, t+T'$ and for all h_{2s} with initial segment h_t',

$$\left\|\hat{\sigma}_1(h_s) - \tilde{E}[\tilde{\sigma}_{1s}|h_s]\right\| < \varepsilon_w, \tag{A.4}$$

then for some $\sigma_2' \in BR(\hat{\sigma}_1; w(h_t'))$, the continuation strategy of σ_2 after the history h_t' agrees with σ_2' for the next T periods.

By assumption, $\exists \kappa > 0$ such that if $\sigma_2^j(h_t) > 0$ then $\sigma_2^j(h_t) > \kappa$. Recall that $\gamma \equiv \min_{y,ij} \rho_{ij}^y$ and set $\psi_w = \frac{1}{2} \min \left\{ \varepsilon_w, (\kappa\gamma)^{T'} \right\}$. Suppose (A.3) holds with this ψ_w. We claim that (A.4) holds for $s = t, \ldots, t + T'$ and for all h_s with initial segment h_t'. Suppose not. The assumption on σ_2 implies that the probability of the continuation history h_s, conditional on the history h_t', is at least $(\kappa\gamma)^{T'}$. Thus,

$$P \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_s] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_s] \right\| \geq \psi_w \middle| h_t \right\} \geq (\kappa\gamma)^{T'},$$

contradicting (A.3), since $(\kappa\gamma)^{T'} > \psi_w$. *Q.E.D.*

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, 58, 1041–1063.

Benabou, R., and G. Laroque (1992): “Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility,” *Quarterly Journal of Economics*, 107, 921–958.

Celentani, M., D. Fudenberg, D. K. Levine, and W. Pesendorfer (1996): “Maintaining a Reputation Against a Long-Lived Opponent,” *Econometrica*, 64(3), 691–704.

Chung, K. L. (1974): *A Course in Probability Theory.* Academic Press, New York.

Cole, H. L., J. Dow, and W. B. English (1995): “Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt,” *International Economic Review*, 36, 365–385.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2003): “Disappearing Private Reputations,” University of Pennsylvania.

Cripps, M. W., and J. P. Thomas (1995): “Reputation and Commitment in Two-Person Repeated Games Without Discounting,” *Econometrica*, 63(6), 1401–1419.

Fudenberg, D., and D. K. Levine (1983): “Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games,” *Journal of Economic Theory*, 31, 251–268.

——— (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4), 759–778.

——— (1992): "Maintaining a Reputation When Strategies Are Imperfectly Observed," *Review of Economic Studies*, 59, 561–579.

——— (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62(1), 103–135.

Fudenberg, D., D. K. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1040.

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

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press, Cambridge, MA.

Hart, S. (1985): "Non-Zero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research*, 10(1), 117–153.

Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–182.

Jackson, M. O., and E. Kalai (1999): "Reputation versus Social Learning," *Journal of Economic Theory*, 88(1), 40–59.

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

Kalai, E., and E. Lehrer (1995): "Subjective Games and Equilibria," *Games and Economic Behavior*, 8(1), 123–163.

Kreps, D., P. R. Milgrom, D. J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27, 245–252.

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

Mailath, G. J., and L. Samuelson (2001): "Who Wants a Good Reputation?," *Review of Economic Studies*, 68(2), 415–441.

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

Phelan, C. (2001): "Public Trust and Government Betrayal," Research Department Staff Report 283, Federal Reserve Bank of Minneapolis.

Sorin, S. (1999): "Merging, Reputation, and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29(1/2), 274–308.