

Imperfect Monitoring and Impermanent Reputations*

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

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

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

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* eventually 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 Game, Commitment.

*We thank the audiences at many seminars for their helpful comments and the National Science Foundation (under grants SES-0095768 and SES-9911219) for financial support.

Contents

1	Introduction	1
2	Illustrative Example and Related Literature	4
3	The Model	7
3.1	The Complete-Information Game	7
3.2	Never an Equilibrium Strategy in the Long Run	9
3.3	The Incomplete-Information Game	12
4	Impermanent Reputations	13
4.1	The “Easy” Case: Player 2’s Beliefs Known	14
4.2	The Harder Case: Player 2’s Beliefs Unknown	16
4.3	Asymptotic Equilibrium Play	17
4.4	Impermanent Restrictions on Behavior	20
5	Proofs of Theorems 1 and 2	21
5.1	Preliminary results	21
5.1.1	Player 2’s Posterior Beliefs	21
5.1.2	Player 2’s Behavior	24
5.1.3	Beliefs about Player 2’s Beliefs	24
5.2	Proof of Theorem 1 (the “Easy” Case)	26
5.3	Proof of Theorem 2 (the Harder Case)	27
6	Imperfect Private Monitoring	35
7	Many Types	37
8	Two Long-Run Players	39
A	Appendix	40
A.1	Proof of Theorem 3 (Section 4.3)	40
A.2	Proof of Theorem 4 (Section 4.4)	42
A.3	Verification of (16) (Section 5.1.3)	44

1 Introduction

The adverse selection approach to reputations is central to the study of long-run relationships. In the complete-information finitely-repeated prisoners' dilemma or chain store game, for example, the intuitively obvious outcome is inconsistent with equilibrium. However, if some player's characteristics are not common knowledge, that player may acquire a reputation for cooperating or playing "tough," rendering the intuitive outcome consistent with equilibrium (Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982), and Milgrom and Roberts (1982)). In other situations, reputation effects impose intuitive limits on the set of equilibria by imposing "high" lower bounds on equilibrium payoffs (Fudenberg and Levine (1989, 1992)).

We explore the long-run possibilities for reputation effects in imperfect monitoring games with a long-lived player facing a sequence of short-lived players. The "short-run" reputation effects in these games are relatively clear-cut. In the absence of incomplete information about the long-lived player, there are many equilibria and the long-lived player's equilibrium payoff is bounded below that player's Stackelberg payoff. However, when there is incomplete information about the long-lived player's type and the latter is patient, reputation effects imply that in every Nash equilibrium, the long-lived player's expected average payoff is arbitrarily close to her Stackelberg payoff.

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 strategy in a game with imperfect monitoring *only* if that strategy eventually plays an equilibrium of the corresponding complete-information game.

More precisely, a commitment type is a long-lived player who plays an exogenously specified strategy. In the incomplete-information game, the long-lived player is either a commitment type or a normal type who maximizes expected payoffs. We show (under some mild conditions) that if the commitment strategy is *not* an equilibrium strategy for the normal type in the complete-information game, then in any Nash equilibrium of the incomplete-information game, if the long-lived player is normal, almost surely the short-lived players will learn that the long-lived player is normal (Theorems 1 and 2). 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. It is straightfor-

ward 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, all public histories are on the equilibrium path. Deviations neither reveal the deviator's type nor trigger punishments. 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 (in terms of altered beliefs) 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 game corresponding to the long-lived player's type.

The intuition of the previous paragraph is relatively straightforward to formalize when the short-lived player's beliefs are known by the long-lived player. Such a case arises, for example, when the short-lived players' actions are public and it is only the long-lived player's actions that are imperfectly monitored. More generally, this case requires that the updating about the long-lived player's actions be independent of the actions taken by the short-lived player. The long-lived player then knows when the short-lived players' beliefs have converged, making deviations from a non-equilibrium commitment strategy irresistibly costless. Our Theorem 1 covers this case.

The situation is more complicated when the short-lived players' beliefs are *not* known by the long-lived player. Now, a player trying to maintain a reputation may not know when her opponent's priors have converged and hence when deviations from the commitment strategy are relatively costless. Making the leap from the preceding intuition to our main result requires showing that there is a set of states of the world under which the short-lived player is relatively certain (in the long run) he faces the commitment type of behavior from the long-lived player, and then using this to show that in the even longer run the long-lived player will (correctly) think that the short-lived player is best responding to the commitment type. This ensures that the long-lived player will deviate from any nonequilibrium commitment strategy, yielding the result (our Theorem 2). This situation is important because the analysis also applies to *games with private monitoring*, situations where there may be *no* public information (Section 6).

The impermanence of reputation arises at the behavioral as well as at

the belief level. Not only do the short-lived players learn that the long-lived player is normal, but asymptotically, continuation play in every Nash equilibrium is a correlated equilibrium of the complete-information game (Theorem 3). Moreover, while the explicit construction of equilibria in reputation games is difficult, we are able to provide a partial converse when the short-lived players' beliefs are known by the long-lived player: Fix a strict Nash equilibrium of the stage game and $\varepsilon > 0$. For all parameter values, there is a Nash equilibrium of the incomplete-information game such that when the long-lived player is normal, with probability at least $1 - \varepsilon$, eventually the stage-game Nash equilibrium is played in every period (Theorem 4). Note that this is true even if the long-lived player is sufficiently patient that reputation effects imply that, in all Nash equilibria of the incomplete-information game, the normal type's expected average payoff is strictly larger than the payoff of that fixed stage-game Nash equilibrium.

For expositional clarity, most of the paper considers a long-lived player, who can be one of two possible types—a commitment and a normal type—facing a sequence of short-lived players. However, most of our results continue to hold when there are many possible commitment types and when the uninformed player is long-lived (Sections 7 and 8).

While the short-run properties of equilibria are interesting, we believe that the long run equilibrium properties are particularly relevant in many situations. For example, the analyst may not know the age of the relationship to which the model is to be applied. Long-run equilibrium properties may then be an important guide to behavior. In other situations, one might take the view of a social planner who is concerned with the continuation payoffs of the long-run player and the fate of all short-run players, even those in the distant future.

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. One attractive mechanism, used in Holmström (1999), Cole, Dow, and English (1995), Mailath and Samuelson (2001), and Phelan (2001), assumes that the type of the long-lived player is itself determined by some stochastic process, rather than being determined once and for all at the beginning of the game. In such a situation, as these papers show, reputations can indeed have long-run implications.

The next section presents a simple motivating example and discusses related literature. Section 3 describes our model. Section 4 presents the statements of the theorems described above. Section 5.1 presents some preliminary results. Theorem 1 is proved in Section 5.2. Our main result, Theorem 2, is proved in Section 5.3.

2 Illustrative Example and Related Literature

Consider an infinite-horizon game involving an infinitely-lived player 1 with discount factor δ and a succession of short-lived player 2's who each live for one period. In each period, the stage game given by

$$\begin{array}{cc}
 & \begin{array}{cc} & 2 \\ & L \quad R \end{array} \\
 \begin{array}{c} 1 \\ \\ \\ \end{array} & \begin{array}{cc} T & \begin{array}{|c|c|} \hline 2, 3 & 0, 2 \\ \hline \end{array} \\ B & \begin{array}{|c|c|} \hline 3, 0 & 1, 1 \\ \hline \end{array} \end{array}
 \end{array} \tag{1}$$

is played. This stage game has a unique Nash equilibrium, BR , which is strict.

If the repeated game has perfect monitoring, then we have a version of the folk theorem: the interval $[1, 2]$ is the set of average discounted subgame-perfect equilibrium payoffs for player 1 as $\delta \rightarrow 1$ (Fudenberg, Kreps, and Maskin (1990)).

In the incomplete-information game, there is a probability $p_0 > 0$ that player 1 is the pure Stackelberg type, i.e., a *commitment* type who plays T in every period, irrespective of history. With complementary probability, player 1 is *normal*, and hence has payoffs in each period given by (1). Fudenberg and Levine (1989) show that for any payoff $u < 2$, for δ sufficiently close to 1, in every Nash equilibrium, the average discounted payoff to player 1 is at least u .¹ An equilibrium (for large δ) with an average payoff of 2 is easy to describe: the normal type begins with T , and plays T in every subsequent period (just like the commitment type). If the normal type ever deviates to B , then she plays B forever thereafter. Player 2 begins with L , plays L as long as T has been played, and plays R forever after B is played. Note that, since there is perfect monitoring, player 2 knows that if he observes B , it must be the case that player 1 is normal, so that BR forever is an equilibrium outcome of the continuation game.

We now describe the game with imperfect monitoring. There are two possible public signals, y' and y'' , which depend on player 1's action i according to the distribution

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

¹Fudenberg and Levine (1989) observe that the argument in that paper cannot be extended to a mixed commitment type, such as a type that plays T with probability 3/4 and B with probability 1/4. Since B is then effectively a noisy signal of the commitment type, the game is more like a game with imperfect monitoring, the case covered by Fudenberg and Levine (1992).

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 are given in the following matrix,

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

The ex ante 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 and player 2's payoffs is the same.

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))).² For our example, it is straightforward to apply the methodology of Abreu, Pearce, and Stacchetti (1990) to show that if $2p > 1 + 2q$, then the set of Nash equilibrium payoffs for large δ is given by the interval

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

There is a continuum of particularly simple equilibria, with player 1 randomizing in every period, with equal probability on T and on B , irrespective of history and player 2's strategy having one period memory. After the signal y' , player 2 plays L with probability α' and R with probability $1 - \alpha'$. After y'' , player 2 plays L with probability α'' and R with probability $1 - \alpha''$, with

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

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

$$\alpha'' = 1 - \frac{1}{2\delta(p-q)}.$$

As in the case of perfect monitoring, 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. Fudenberg and Levine (1992) show that in this case as well, for *any* payoff $u < 2$, there is δ sufficiently close to 1 such that

²In other words, the folk theorem of Fudenberg, Levine, and Maskin (1994) does not hold when there are short-lived players.

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). Establishing that u is a lower bound on equilibrium payoffs is significantly complicated by the imperfect monitoring. Player 2 no longer observes the action choices of player 1, and so any attempt by player 1 to manipulate the beliefs of player 2 must be mediated through the full support signals. The explicit construction of equilibria in this case is also very difficult.

In this example, player 2's posterior belief that player 1 is the Stackelberg type is independent of 2's actions and hence public information. As such the example falls within the coverage of our Theorem 1. To develop intuition, temporarily restrict attention to 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 is the prior, and continuation play is also 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 informative, and so almost surely, when player 1 is normal, beliefs must converge to zero probability on the Stackelberg type.³

Our analysis is complicated by the fact that we do not restrict attention to Markov perfect equilibria, as well as the possibility of more complicated commitment types than the pure Stackelberg type (for example, we allow for nonstationary history-dependent mixing). In particular, uninformative signals may have future ramifications.

While some of our arguments and results are reminiscent of the recent literature on rational learning and merging, there are also 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 one-shot Nash equilibrium of the complete-information 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, and the information processing in our model is complicated by the imperfect monitoring.

³Benabou 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.

A key idea in our results (in particular, Lemmas 2 and 5) is that if signals are statistically informative about a player’s behavior, then there is little *asymptotic* value to other players learning private information that has a nontrivial asymptotic impact on the first player’s behavior. Similar ideas play an important role in merging arguments, which provide conditions under which a true stochastic process and beliefs over that process converge. Sorin (1999), for example, unifies much of the existing reputation literature as well as recent results on repeated games with incomplete information using merging. Kalai and Lehrer (1995), again using merging, provide a simple argument that in reputation games, asymptotic continuation play is a subjective correlated equilibrium of the complete-information game (that result is immediate in our context, since we begin with a Nash equilibrium of the incomplete-information game, while it is a harder for Kalai and Lehrer (1995) since their hypothesis is a weaker). Subjective correlated equilibrium is a significantly weaker solution concept than objective correlated equilibrium. We discuss the relationship in Section 4.3, where we show that asymptotic continuation play is an objective correlated equilibrium of the complete-information game.

The idea that reputations are temporary is a central theme of Jackson and Kalai (1999), who are interested in reputation in finitely repeated normal form games (for which Fudenberg and Maskin (1986) prove a reputation folk theorem). Jackson and Kalai (1999) prove that if the finitely repeated normal-form game is itself repeated (call the finite repetition of the original stage a *round*), with new players (although not a new draw from a rich set of types) in each round, then eventually, reputations cannot affect play in the finitely repeated game. While the conclusion looks the same as ours, 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. Moreover, 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 $i \in \{1, 2, \dots, I\} \equiv I$ and player 2 simultaneously chooses an action $j \in \{1, 2, \dots, J\} \equiv J$. 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 (respectively, 2) from the action i (resp., j) and signal y is given by $f_1(i, y)$ (resp., $f_2(j, y)$). The ex ante stage game payoffs are $\pi_1(i, j) = \sum_y f_1(i, y) \rho_{ij}^y$ and $\pi_2(i, j) = \sum_y f_2(j, y) \rho_{ij}^y$.

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.

Players only observe the realizations of the public signal and their own past actions (the period- t player 2 knows the action choices of the previous player 2’s). Player 1 in period t has a *private history*, consisting of the public signals and her own past actions, denoted by $h_{1t} \equiv ((i_0, y_0), (i_1, y_1), \dots, (i_{t-1}, y_{t-1})) \in H_{1t} \equiv (I \times Y)^t$. Similarly, a *private history* for player 2 is denoted $h_{2t} \equiv ((j_0, y_0), (j_1, y_1), \dots, (j_{t-1}, y_{t-1})) \in H_{2t} \equiv (J \times Y)^t$. Let $\{\mathcal{H}_{\ell t}\}_{t=0}^{\infty}$ denote the filtration on $(I \times J \times Y)^{\infty}$ induced by the private histories of player $\ell = 1, 2$. The *public history* observed by both players is the sequence $(y_0, y_1, \dots, y_{t-1}) \in Y^t$. 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 describe circumstances under which this assumption can be weakened in Section 4.1. 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 $j \in J$, there are I linearly independent columns in the matrix $(\rho_{ij}^y)_{y \in Y, i \in I}$.

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

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 $\mathcal{H}_{\ell t}$.⁴

⁴This expectation is well-defined, since I, J , and Y are finite.

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 ,⁵

$$E^\sigma[\pi_2(i_t, j_t) \mid \mathcal{H}_{2t}] \geq E^\sigma[\pi_2(i_t, j) \mid \mathcal{H}_{2t}], \quad \forall j \in J \text{ } P^\sigma\text{-a.s.}$$

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 :*

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

This requires that after any history that can arise with positive probability under the equilibrium profile, player 1's strategy maximize her continuation expected utility. Hence, if σ^* is a Nash equilibrium, then for all σ_1 and for all t , P^{σ^*} -almost surely

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

Since every public history occurs with positive probability, the outcome of any Nash equilibrium is a perfect Bayesian equilibrium outcome.

3.2 Never an Equilibrium Strategy in the Long Run

Suppose $(\bar{\sigma}_1, \bar{\sigma}_2)$ is an equilibrium of the complete-information game and that we extend this to an incomplete-information game by introducing the possibility of a commitment type who plays $\bar{\sigma}_1$. The profile in which player 1 always plays like the commitment type and player 2 follows $\bar{\sigma}_2$ is then an equilibrium of the incomplete-information game. Moreover, player 2 learns nothing about the type of player 1 in this equilibrium. Hence, player 1 can maintain a permanent reputation for behavior that would be an equilibrium without that reputation, i.e., in the complete-information game.

⁵Note that j , the action of player 2, on the right hand side of the inequality is not random, and so the right expectation is being taken only with respect to i_t , the choice of player 1.

More generally, there may be no difficulty in maintaining a reputation for behavior that features nonequilibrium play in the first ten periods and thereafter switches to $\bar{\sigma}_1$. Questions of whether player 2 will learn player 1's type can only be settled by long-run characteristics of strategies, independent of initial histories. We accordingly introduce the concept of a strategy's being *never an equilibrium strategy in the long run*. Such a strategy for player 1 has the property that for all best responses by player 2 and all histories h_{1t} , there is always a profitable deviation for player 1 in periods beyond some sufficiently large T . We emphasize that in the following definition, the $BR(\bar{\sigma}_1)$ is the set of player 2 best responses in the complete-information game.

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

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

It is possible for a strategy to never be an equilibrium strategy in the long run for some discount factors, but not for others.

This definition is most easily interpreted when the strategy is either simple or implementable by a finite automaton:

Definition 3 (1) *A behavior strategy σ_1 is public if it is measurable with respect to $\{\mathcal{H}_t\}$, so that the mixture over actions induced in each period depends only upon the public history.*

(2) *A behavior strategy σ_1 is simple if it is a constant function, i.e., induces the same (possibly degenerate) mixture over Δ^I after every history.*

(3) *A public strategy σ_1 is implementable by a finite automaton if there exists a finite set W , an action function $d : W \rightarrow \Delta^I$, a transition function $\varphi : W \times Y \rightarrow W$, 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 .*

Any pure strategy is realization equivalent to a public strategy. A simple strategy is clearly public and is implementable by a finite automaton with a single state.

The following Lemma shows that for simple strategies or strategies implementable by a finite automata, being never an equilibrium in the long run

is essentially equivalent to not being part of a Nash equilibrium of the stage game or the complete-information repeated game. The point of Definition 2 is to extend this concept to strategies that have transient initial phases or that never exhibit a stationary structure.

Lemma 1 *Assume the monitoring has full support (Assumption 1).*

1.1 *Suppose player 2 has a unique best reply to some mixture $\varsigma \in \Delta^I$. The simple strategy of always playing ς is never an equilibrium strategy in the long run if and only if ς is not part of a stage-game Nash equilibrium.*

1.2 *Suppose $\bar{\sigma}_1$ is a public strategy 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.*

Proof. We prove only part 2, since part 1 is similar (but easier). The only if direction 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. ■

The example from Section 2 illustrates the necessity of the condition in Lemma 1 that player 2 have a unique best response. The simple strategy that places equal probability on T and B is part of many equilibria of the complete-information game (as long as $\delta > 1/[2(p-q)]$), and hence fails the criterion for being never an equilibrium strategy in the long run. However, this strategy is not part of an equilibrium of the stage game, in contrast to Lemma 1.1. 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 simple strategy that always plays such a mixture is not part of a stage-game equilibrium and is never an equilibrium strategy in the long run.

If we were only interested in the presence of “Stackelberg” commitment types, and the attendant lower bounds on player 1’s ex ante payoffs, it would suffice to consider commitment types who follow simple strategies. However, allowing more general types leaves the structure of the argument unaffected while simplifying our discussion of the case where player 2 is also long-lived (see Section 8).

3.3 The Incomplete-Information Game

We now formally describe the game with incomplete information about the type of player 1. For expositional clarity only, much of our analysis focuses on the case of one commitment type. Section 7 discusses the case of many commitment types.

At time $t = -1$ a type of player 1 is selected. With probability $1 - p_0$ she is the “normal” type, denoted n , with the preferences described above. With probability $p_0 > 0$ she is a “commitment” type, denoted c , who plays a fixed, repeated-game strategy $\hat{\sigma}_1$.

A state of the world is now a type for player 1 and sequence of actions and signals. The set of states is then $\Omega = \{n, c\} \times (I \times J \times Y)^\infty$. 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 an uninformed player expects play to evolve. The strategy profile $\hat{\sigma} = (\hat{\sigma}_1, \sigma_2)$ (respectively $\tilde{\sigma}$) determines a probability measure \hat{P} (respectively \tilde{P}) over Ω , which describes how play evolves when player 1 is the commitment (respectively 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 $\{\mathcal{H}_{1t}\}_{t=0}^\infty$, $\{\mathcal{H}_{2t}\}_{t=0}^\infty$, 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). As usual, denote by $\mathcal{H}_{\ell\infty}$ the σ -algebra generated by $\cup_{t=0}^\infty \mathcal{H}_{\ell t}$.

For any repeated-game behavior strategy $\sigma_1 : \cup_{t=0}^\infty H_{1t} \rightarrow \Delta^I$, denote by σ_{1t} the t th period behavior strategy, so that σ_1 can be viewed as the sequence of functions $(\sigma_{10}, \sigma_{11}, \sigma_{12}, \dots)$ with $\sigma_{1t} : H_{1t} \rightarrow \Delta^I$. 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 func-

tion as in the complete-information game. Player 2, on the other hand, is maximizing $E[\pi_2(i_t, j) | \mathcal{H}_{2t}]$, so that after any history h_{2t} , he is updating his beliefs over the type of player 1 that he is facing.⁶ 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}_{2t} -measurable random variable $p_t : \Omega \rightarrow [0, 1]$. By Assumption 1, Bayes' rule determines this posterior after all sequences of signals. Thus, in period t , player 2 is maximizing

$$p_t \hat{E}[\pi_2(i_t, j) | \mathcal{H}_{2t}] + (1 - p_t) \tilde{E}[\pi_2(i_t, j) | \mathcal{H}_{2t}]$$

P -almost surely. At any Nash equilibrium of this game, the belief p_t is a bounded martingale with respect to the filtration $\{\mathcal{H}_{2t}\}_t$ and measure P .⁷ 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 \tilde{P} -submartingale and a \hat{P} -supermartingale with respect to the filtration $\{\mathcal{H}_{2t}\}$.

A final word on notation: The expression $\tilde{E}[\sigma_{1t} | \mathcal{H}_{2s}]$ is the standard conditional expectation, viewed as a \mathcal{H}_{2s} measurable random variable on Ω , while $\hat{E}[\sigma_1(h_{1t}) | h_{2s}]$ is the conditional expected value of $\sigma_1(h_{1t})$ (with h_{1t} viewed as a random history) conditional on the observation of the history h_{2s} .

4 Impermanent Reputations

Consider an incomplete-information game, with a commitment type strategy that is never an equilibrium strategy in the long run. Suppose there is a Nash equilibrium 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 can ascertain which type he is facing), and hence must believe, on this set of histories, that the normal and commitment types are playing the same strategies on average. But then player 2 must eventually, again on the same set of histories, best reply to the average behavior of the commitment type. Since the commitment type's

⁶As in footnote 5, j is not random and the expectation is being taken with respect to i_t , the action choice of player 1.

⁷These 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).

behavior is never an equilibrium strategy in the long run, player 1 does not find it optimal to play the commitment-type strategy in response to 2's best response, leading us very close to a contradiction.

There are two difficulties in making this argument precise. First, since the game has imperfect monitoring, player 2 has imperfect knowledge of player 1's private history and thus the continuation strategy of player 1. If the commitment type strategy is not pure, it may be that the normal type is following a private strategy that on average is like the public commitment strategy, but which is different from the latter on every history. Second, the set of histories upon which the argument proceeded is typically not known by either player at any point (although it will be in $\mathcal{H}_{2\infty}$). Consequently, player 1 may never know that player 2 is best responding to the average play of the commitment type.

These two difficulties interact. Our first and easier result (Theorem 1) is for the case where the informativeness of the signal about player 1's action is independent of player 2's action. In this case, in any equilibrium, player 2's beliefs about the type of player 1 are public and so the second difficulty does not arise. We can then surmount the first difficulty to show that reputations are impermanent, even when the commitment type is following a mixed public strategy.

The second, harder result (Theorem 2) is for the case where the informativeness of the signal about player 1's action depends on player 2's action. This case is important because, as we describe in Section 6, it also covers games with private monitoring. In this case, we can only show that reputations are impermanent when we remove the first difficulty by imposing the stronger requirement that the commitment type is following a pure (though not necessarily simple or finitely-implementable) strategy.

Theorem 1 is presented in the next subsection with its proof given in Section 5.2, Theorem 2 is presented in Section 4.2 and proved in Section 5.3. (Some preliminaries are presented in Section 5.1.) The behavioral implications of the theorems are discussed in Sections 4.3 and 4.4.

4.1 The “Easy” Case: Player 2's Beliefs Known

In the “easy” case, player 2's beliefs about player 1 are known by player 1.

The full support assumption (Assumption 1) implies that player 1 in general does not know the action choice of player 2. Under the following assumption, however, player 1 can calculate 2's inference without knowing 2's action:

Assumption 3 (Independence) For any (possibly mixed) actions $\zeta_1 \in \Delta^I$, signal $y \in Y$, and actions i, j , and j' ,

$$\Pr\{i|y, j, \zeta_1\} = \Pr\{i|y, j', \zeta_1\},$$

where $\Pr\{i|y, j, \zeta_1\}$ is the posterior probability of player 1 having chosen pure action i , given mixed action ζ_1 and given that player 2 observed signal y after playing action j .

Theorem 1 Suppose ρ satisfies Assumptions 1, 2, and 3. Suppose $\hat{\sigma}_1$ is a public strategy with finite range that is never an equilibrium strategy in the long run. Then in any Nash equilibrium, $p_t \rightarrow 0$ \tilde{P} -almost surely.

The definition of never an equilibrium in the long run requires player 1's period- t deviation to generate an expected payoff increase, conditional on reaching period t , of at least ε . Our proof rests on the argument that if players are eventually almost certain that the normal type player 1 is behaving like a commitment type that is never an equilibrium in the long run, then the normal type will have a profitable deviation. Without the ε wedge in this definition, it is conceivable that while players become increasingly certain that the normal type is playing like the commitment type, the payoff premium to deviating from the nonequilibrium commitment-type strategy declines sufficiently rapidly as to ensure that the players are never certain enough to support a deviation. The ε -uniform bound on the profitability of a deviation precludes this possibility.

Section 5.1.2 explains how the assumptions that $\hat{\sigma}$ has a finite range plays a role similar to that of the ε just described. The requirement that $\hat{\sigma}_1$ be public ensures that whenever player 2 is convinced that player 1 is playing like the commitment type, player 2 can identify the period- t strategy realization $\hat{\sigma}_1(h_{1t})$ and play a best response.

A sufficient condition for Assumption 3 is that the public signal y be 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

$$\rho_{ij}^y = \rho_i^{y_1} \rho_j^{y_2} \quad \forall i, j, y. \quad (3)$$

The full-support Assumption 1 can be relaxed if (3) holds. The key ingredient in the proof of Theorem 1 is that players 1 and 2 are symmetrically informed about 2's beliefs, and that the signal not reveal player 1's action (so

that trigger profiles are not equilibria). Assumption 1 can thus be replaced by the requirement that, for all i and $y_1 \in Y_1$,

$$\rho_i^{y_1} > 0.$$

Assumption 2, in the presence of (3), is equivalent to the requirement that there are I linearly independent columns in the matrix

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

Since the key implication of Assumption 3 is that player 1 knows player 2's posterior belief, an alternative to Assumption 3 is to assume that player 2's *actions* are public, while maintaining imperfect public monitoring of player 1's actions. In this case, $Y = Y_1 \times J$, where Y_1 is the set of public signals of player 1's actions, and

$$\rho_{ij}^{(y_1, j')} = 0 \tag{4}$$

for all $i \in I, j \neq j' \in J$, and $y_1 \in Y$. The public nature of player 2's actions implies that $\mathcal{H}_{2t} = \mathcal{H}_t$, and hence p_t is measurable with respect to \mathcal{H}_t (and so player 1 knows the posterior of belief of player 2).

When player 2's actions are public, the full support assumption is

$$\rho_{ij}^{(y_1, j)} > 0$$

for all $(i, j) \in I \times J$ and $y_1 \in Y_1$, while the identification assumption is now that for all $j \in J$, there are I linearly independent columns in the matrix

$$\left(\rho_{ij}^{(y_1, j)} \right)_{y_1 \in Y_1, i \in I}.$$

4.2 The Harder Case: Player 2's Beliefs Unknown

The harder case is where player 2's beliefs about player 1 are not known by player 1. Our method of proof requires that player 1 can draw inferences about player 2's actions, and the following assumption allows this:

Assumption 4 *For all $i \in I$, there are J linearly independent columns in the matrix $(\rho_{ij}^y)_{y \in Y, j \in J}$.*

This assumption is dramatically weaker than Assumption 3. Consider the example of Section 2, except that player 2's choice of L or R is private. Let ρ'_{ij} be the probability of the signal y' under the action profile $ij \in$

$\{T, B\} \times \{L, R\}$, so that the probability of the signal y'' is given by $1 - \rho'_{ij}$. Assumption 2 requires $\rho'_{TL} \neq \rho'_{BL}$ and $\rho'_{TR} \neq \rho'_{BR}$, while Assumption 4 requires $\rho'_{TL} \neq \rho'_{TR}$ and $\rho'_{BL} \neq \rho'_{BR}$. The assumptions are satisfied if $\rho'_{TL} > \rho'_{BL}$ and $\rho'_{TR} < \rho'_{BR}$, so that y' is a signal that player 1 has played T if player 2 had played L , but is a signal that she had played B if 2 had played R . Unless player 1 knows the action of player 2, she will not know how the signal is interpreted.

As we discussed at the beginning of this section, the cost of weakening Assumption 3 to Assumption 4 is that we must assume the commitment type does not randomize. The commitment type's strategy, while pure, can still depend upon histories in arbitrarily complicated ways. We also emphasize that we are *not* imposing any restrictions on the normal type's behavior (other than it be a best response to the behavior of the short-lived players).

Theorem 2 *Suppose ρ satisfies Assumptions 1, 2, and 4. Suppose $\hat{\sigma}_1$ is a pure strategy that is never an equilibrium strategy in the long run. Then in any Nash equilibrium, $p_t \rightarrow 0$ \tilde{P} -almost surely.*

Since any pure strategy is realization equivalent to a public strategy, it is again the case that whenever player 2 is convinced that player 1 is playing like the commitment type, player 2 can identify the period- t strategy realization $\hat{\sigma}_1(h_{1t})$ and play a best response.

4.3 Asymptotic Equilibrium Play

We now explore the implications for equilibrium play of the impermanence of reputations. More precisely, we will show that in the limit, the normal type of player 1 and player 2 play a correlated equilibrium of the complete-information game. Hence, differences in the players' subjective beliefs about how play will continue vanish in the limit. This strengthens the result on convergence to *subjective* equilibria (see below) obtained by Kalai and Lehrer (1995, Corollary 4.4.1). To begin, we describe some notation for the correlated equilibrium of our repeated games with imperfect monitoring.

We use the term *continuation game* for the game with initial period in period t , ignoring the period t -histories.⁸ We use the notation $t' = 0, 1, 2, \dots$ for a period of play in a continuation game (which may be the original game)

⁸Since a strategy profile of the original game induces a probability distribution over t -period histories, $H_{1t} \times H_{2t}$, we can view the period t continuation, together with a type space $H_{1t} \times H_{2t}$ and induced distribution on that type space, as a Bayesian game. Different strategy profiles in the original game induce different distributions over the type space in the continuation 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 $s_{1t'} : H_{1t'} \rightarrow I$ for $t' = 0, 1, \dots$ ⁹ Thus, $s_{1t'} \in I^{H_{1t'}}$ and $s_1 \in I^{\cup_{t'} H_{1t'}} \equiv S_1$, and similarly $s_2 \in S_2 \equiv J^{\cup_{t'} H_{2t'}}$. The spaces S_1 and S_2 are countable products of finite sets. We equip the product space $S \equiv S_1 \times S_2$ with the σ -algebra generated by the cylinder sets, denoted by \mathcal{S} . Denote the players' payoffs in the infinitely repeated game (as a function of these pure strategies) as follows

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

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.

In the following definitions, we follow Hart and Schmeidler (1989) in using the ex ante definition of correlated equilibria for infinite pure strategy sets. Note also that we need player 2's incentive compatibility conditions to hold at all times t' , because of player 2's zero discounting.

Definition 4 *A subjective correlated equilibrium of the complete-information game is a pair of measures μ_ℓ , $\ell = 1, 2$, on (S, \mathcal{S}) such that for all \mathcal{S} -measurable functions $\zeta_1 : S_1 \rightarrow S_1$, $\zeta_2 : S_2 \rightarrow S_2$*

$$\int_S [u_1(s_1, s_2) - u_1(\zeta_1(s_1), s_2)] d\mu_1 \geq 0; \quad (5)$$

$$\int_S [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \zeta_2(s_2))] d\mu_2 \geq 0, \quad \forall t'. \quad (6)$$

A correlated equilibrium of the complete-information game is a subjective correlated equilibrium satisfying $\mu_1 = \mu_2$.

Let \mathcal{M} denote the space of probability measures μ on (S, \mathcal{S}) , equipped with the product topology. Then, a sequence μ_n converges to μ if, for each $\tau > 0$, we have

$$\mu_n|_{I^{(I \times Y)^\tau} \times J^{(J \times Y)^\tau}} \rightarrow \mu|_{I^{(I \times Y)^\tau} \times J^{(J \times Y)^\tau}}. \quad (7)$$

Moreover, \mathcal{M} is sequentially compact with this topology. Payoffs for players 1 and 2 are extended to \mathcal{M} in the obvious way. Since player 1's payoffs are

⁹Note that we have used σ_1 for general behavior strategies, not only pure strategies.

discounted, the product topology is strong enough to guarantee continuity of $u_1 : \mathcal{M} \rightarrow \mathbb{R}$. Each player 2's payoff is trivially continuous. The set of mixed strategies for player ℓ is denoted by \mathcal{M}_ℓ .

Fix an equilibrium of the incomplete-information game with imperfect monitoring. When player 1 is the normal (respectively, commitment) type, the monitoring technology and the behavior strategies $(\tilde{\sigma}_1, \sigma_2)$ (resp., $(\hat{\sigma}_1, \sigma_2)$) induce a probability measure $\tilde{\phi}_t$ (resp., $\hat{\phi}_t$) on the t -period histories $(h_{1t}, h_{2t}) \in H_{1t} \times H_{2t}$. 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 $\tilde{\lambda}^{h_{1t}} \in \mathcal{M}_1$ for the continuation game. Similarly, the commitment type will play a mixed strategy $\hat{\lambda}^{h_{1t}} \in \mathcal{M}_1$ for the continuation game and player 2 will form his posterior $p_t(h_{2t})$ and play the mixed strategy $\lambda^{h_{2t}} \in \mathcal{M}_2$ in the continuation game. Conditional on player 1 being normal, the composition of the probability measure $\tilde{\phi}_t$ and the measures $(\tilde{\lambda}^{h_{1t}}, \lambda^{h_{2t}})$ induces a joint probability measure, $\tilde{\rho}_t$, on the pure strategies in the continuation game and player 2's posterior (the space $S \times [0, 1]$). Similarly, conditional upon player 1 being the commitment type, there is a measure $\hat{\rho}_t$ on $S \times [0, 1]$. Let $\tilde{\mu}_t$ denote the marginal of $\tilde{\rho}_t$ on S and $\hat{\mu}_t$ denote the marginal of $\hat{\rho}_t$ on S .

At the fixed equilibrium, the normal type is playing in an optimal way from time t onwards given her available information. This implies that for all \mathcal{S} -measurable functions $\zeta_1 : S_1 \rightarrow S_1$,

$$\int_S u_1(s_1, s_2) d\tilde{\mu}_t \geq \int_S u_1(\zeta_1(s_1), s_2) d\tilde{\mu}_t. \quad (8)$$

Let $\mathcal{S} \times \mathcal{B}$ denote the product σ -algebra on $S \times [0, 1]$ generated by \mathcal{S} on S and the Borel σ -algebra on $[0, 1]$. Player 2 is also playing optimally from time t onwards, which implies that for all $\mathcal{S} \times \mathcal{B}$ -measurable functions $\xi_2 : S_2 \times [0, 1] \rightarrow S_2$, and for all t' ,

$$\int_{S \times [0, 1]} u_2'(s_1, s_2) d(p_0 \hat{\rho}_t + (1-p_0) \tilde{\rho}_t) \geq \int_{S \times [0, 1]} u_2'(s_1, \xi_2(s_2, p_t)) d(p_0 \hat{\rho}_t + (1-p_0) \tilde{\rho}_t).$$

Comparing the previous two inequalities with (5) and (6), it is clear that the equilibrium behavior from period t onwards is a subjective correlated equilibrium for the continuation game for all t .

If we had metrized \mathcal{M} , a natural formalization of the idea that asymptotically, the normal type and player 2 are playing a correlated equilibrium is that the distance between the set of correlated equilibria and the induced distributions $\tilde{\mu}_t$ on S goes to zero. While \mathcal{M} is metrizable, a simpler and

equivalent formulation is that the limit of every convergent subsequence of $\{\tilde{\mu}_t\}$ is a correlated equilibrium. This equivalence is an implication of the sequential compactness of \mathcal{M} , since every subsequence of $\{\tilde{\mu}_t\}$ has a convergent sub-subsequence. The proof is in Section A.1.

Theorem 3 *Fix a Nash equilibrium of the incomplete-information game and suppose $p_t \rightarrow 0$ \tilde{P} -almost surely. Let $\tilde{\mu}_t$ denote the distribution on S induced in period t by the Nash equilibrium. The limit of every convergent subsequence is a correlated equilibrium of the complete-information game.*

Since players have access to a coordination device, namely histories, in general it is not true that Nash equilibrium play of the incomplete-information game eventually looks like Nash equilibrium play of the complete-information game.

Suppose the Stackelberg payoff is not a correlated 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 3 lies in the order of quantifiers: while Fudenberg and Levine (1992) fix the prior, p_0 , and then select $\bar{\delta}(p_0)$ large (with $\bar{\delta}(p_0) \rightarrow 1$ as $p_0 \rightarrow 0$), we fix δ and examine asymptotic play, so that eventually p_t is sufficiently small that $\bar{\delta}(p_t) > \delta$.

We do not know if Nash equilibrium play in the incomplete-information game eventually looks like a public randomization over Nash equilibrium play in the complete-information game.¹⁰

4.4 Impermanent Restrictions on Behavior

We now provide a partial converse to the previous section by identifying a class of equilibria of the complete-information game to which equilibrium play of the incomplete-information game can converge. The construction of equilibria in incomplete-information games is difficult, and so we restrict

¹⁰As far as we know, it is also not known whether the result of Fudenberg and Levine (1994, Theorem 6.1, part (iii)) extends to correlated equilibrium. That is, for moral hazard mixing games and for large δ , is it true that the long-run player's maximum *correlated* equilibrium payoff is lower than when monitoring is perfect? We believe that, at least for simple games like that described in Section 2, allowing for correlation does not increase the long-lived player's payoff.

attention to the case in which the posterior beliefs of player 2 are known by player 1.

Recall that in the example of Section 2, the stage game has a (unique) strict Nash equilibrium BR . Moreover, this achieves player 1's minmax utility. 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, for much of the initial history of the game, in *every* equilibrium, play is like TR . On the other hand, an implication of Theorem 4 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, reputation effects may ensure that for a long period of time, equilibrium play will be very different from BR (this is just Fudenberg and Levine (1992)). Theorem 4 is consistent with this, since it only claims that in the equilibrium of interest, BR is *eventually* played in every period with high probability. The second issue is that, even if reputation effects are not initially operative (because the initial belief that player 1 is the commitment type is low relative to the discount factor), with positive probability (albeit small), a sequence of signals will arise that will make reputation effects operative (because the posterior that player 1 is the commitment type is increased sufficiently).

Theorem 4 *Suppose the assumptions of Theorem 1 are satisfied (i.e., ρ satisfies Assumptions 1, 2, and 3, and $\hat{\sigma}_1$ is a public strategy with finite range that is never an equilibrium strategy in the long run). Suppose (i^*, j^*) is a strict Nash equilibrium of the stage game. 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 $1 - \varepsilon$.*

The proof is in Section A.2.

5 Proofs of Theorems 1 and 2

5.1 Preliminary results

5.1.1 Player 2's Posterior Beliefs

The first step is to show that either player 2's expectation (given his private history) of the strategy played by the commitment type is in the limit iden-

tical to his expectation of the strategy played by the normal 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, for a given private history for player 2, the distributions generating his 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 for each private history. In the statement of the following Lemma, h_{1t} is to be interpreted as a function from Ω to $(I \times Y)^t$.

Lemma 2 *At any Nash equilibrium of a game satisfying Assumptions 1 and 2,*¹¹

$$\lim_{t \rightarrow \infty} p_t(1 - p_t) \left\| \hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}] \right\| = 0, \quad P\text{-a.s.} \quad (9)$$

Proof. Let $p_{t+1}(h_{2t}; j_t, y_t)$ 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_{2t} . By Bayes' rule,

$$p_{t+1}(h_{2t}; j_t, y_t) = \frac{p_t \Pr[y_t | h_{2t}, j_t, c]}{p_t \Pr[y_t | h_{2t}, j_t, c] + (1 - p_t) \Pr[y_t | h_{2t}, j_t, n]}.$$

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

$$p_{t+1}(h_{2t}; j_t, y_t) = \frac{p_t \sum_{i \in I} \rho_{ij_t}^{y_t} \hat{E}[\hat{\sigma}_1^i(h_{1t}) | h_{2t}]}{\sum_{i \in I} \rho_{ij_t}^{y_t} \left(p_t \hat{E}[\hat{\sigma}_1^i(h_{1t}) | h_{2t}] + (1 - p_t) \tilde{E}[\tilde{\sigma}_1^i(h_{1t}) | h_{2t}] \right)}.$$

Rearranging this formula allows us to write

$$\begin{aligned} p_t(1 - p_t) \sum_{i \in I} \rho_{ij_t}^{y_t} \left(\hat{E}[\hat{\sigma}_1^i(h_{1t}) | h_{2t}] - \tilde{E}[\tilde{\sigma}_1^i(h_{1t}) | h_{2t}] \right) \\ = (p_{t+1} - p_t) \sum_{i \in I} \rho_{ij_t}^{y_t} \left(p_t \hat{E}[\hat{\sigma}_1^i(h_{1t}) | h_{2t}] + (1 - p_t) \tilde{E}[\tilde{\sigma}_1^i(h_{1t}) | h_{2t}] \right). \end{aligned}$$

The summation on the right is bounded above by $\max_i \rho_{ij_t}^{y_t} < 1$, thus

$$p_t(1 - p_t) \left| \sum_{i \in I} \rho_{ij_t}^{y_t} \left(\hat{E}[\hat{\sigma}_1^i(h_{1t}) | h_{2t}] - \tilde{E}[\tilde{\sigma}_1^i(h_{1t}) | h_{2t}] \right) \right| \leq |p_{t+1} - p_t|.$$

¹¹We use $\|x\|$ to denote the sup-norm on \mathbb{R}^n .

For any fixed realization y of the signal y_t , it follows that for all h_{2t} and j_t ,

$$\begin{aligned} p_t(h_{2t})(1 - p_t(h_{2t})) & \left| \sum_{i \in I} \rho_{ij_t}^y \left(\hat{E}[\hat{\sigma}_1^i(h_{1t})|h_{2t}] - \tilde{E}[\tilde{\sigma}_1^i(h_{1t})|h_{2t}] \right) \right| \\ & \leq \max_{y'} |p_{t+1}(h_{2t}; j_t, y') - p_t(h_{2t})|. \end{aligned} \quad (10)$$

Since p_t is a P -almost sure convergent sequence, it is Cauchy P -almost surely.¹² So the right hand side of (10) converges to zero P -almost surely. Thus, for any y ,

$$p_t(1 - p_t) \left| \sum_{i \in I} \rho_{ij_t}^y \left(\hat{E}[\hat{\sigma}_{1t}^i|\mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t}^i|\mathcal{H}_{2t}] \right) \right| \rightarrow 0, \quad P\text{-a.s.} \quad (11)$$

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 (11) implies (9). Let Π_{j_t} be a $|Y| \times |I|$ matrix whose y -th row, for each signal $y \in Y$, contains the terms $\rho_{ij_t}^y$ for $i = 1, \dots, |I|$. Then as (11) holds for all y (and Y is finite), it can be restated as

$$p_t(1 - p_t) \left\| \Pi_{j_t} \left(\hat{E}[\hat{\sigma}_{1t}|\mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_{2t}] \right) \right\| \rightarrow 0, \quad P\text{-a.s.}, \quad (12)$$

where $\|\cdot\|$ 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 (12), we then get

$$\begin{aligned} p_t(1 - p_t) & \left\| \Pi_{j_t} \left(\hat{E}[\hat{\sigma}_{1t}|\mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_{2t}] \right) \right\| \\ & \geq p_t(1 - p_t)b \left\| \hat{E}[\hat{\sigma}_{1t}|\mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_{2t}] \right\| \rightarrow 0, \quad P\text{-a.s.}, \end{aligned}$$

which implies (9). ■

Condition (9) says that either player 2's best prediction of the normal type's behavior at the current stage is identical to his best prediction of the commitment type's behavior (that is, $\|\hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}]\| \rightarrow 0$ P -almost surely), or the type is revealed (that is, $p_t(1 - p_t) \rightarrow 0$ P -almost

¹²Note that the analysis is now *global*, rather than *local*, in that we treat all the expressions as functions on Ω .

surely). However, $\lim p_t < 1$ \tilde{P} -almost surely, and hence (9) implies a simple corollary:¹³

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

$$\lim_{t \rightarrow \infty} p_t \left\| \hat{E}[\hat{\sigma}_{1t} \mid \mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t} \mid \mathcal{H}_{2t}] \right\| = 0, \quad \tilde{P}\text{-a.s.}$$

5.1.2 Player 2's Behavior

We next show that if player 2 believes player 1's strategy is close to the commitment strategy, then 2's best response is a best response to the commitment type.

Lemma 3 *Suppose $\hat{\sigma}_1$ has a finite range. There exists $\psi > 0$ such that for any history h_{1s} and any $\zeta_1 \in \Delta^I$ satisfying $\|\zeta_1 - \hat{\sigma}_1(h_{1s})\| < \psi$, a best response to ζ_1 is also a best response to $\hat{\sigma}_1(h_{1s})$.*

Proof. The best response correspondence is upper semi-continuous. Thus, for any mixed action $\hat{\sigma}_1(h_{1s})$, there exists a $\psi(\hat{\sigma}_1(h_{1s})) > 0$ such that a best response to any mixed action ζ_1 which satisfies $\|\zeta_1 - \hat{\sigma}_1(h_{1s})\| < \psi(\hat{\sigma}_1(h_{1s}))$ is also a best response to $\hat{\sigma}_1(h_{1s})$. We then let ψ be the minimum of such $\psi(\hat{\sigma}_1(h_{1s}))$ over the finite range of $\hat{\sigma}$. ■

Thus, if player 2 believed his opponent was “almost” the commitment type, then 2 would be playing the same equilibrium action as if he was *certain* he was facing the commitment type.

The finiteness of the range of $\hat{\sigma}$ ensures that the minimum of the $\psi(\hat{\sigma}_1(h_{1s}))$ is strictly positive. Otherwise, player 2 could conceivably become ever more convinced that the normal type is playing like the commitment type, only to have the commitment type's stage-game action drift in such a way that player 2 is never sufficiently convinced of commitment-type play to choose a best response to the commitment type.

5.1.3 Beliefs about Player 2's Beliefs

Lemma 4 *Suppose Assumptions 1 and 2 are satisfied. Suppose there exists a \tilde{P} -positive measure set of states A on which $\lim_{t \rightarrow \infty} p_t(\omega) > 0$. Then, for*

¹³Since 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.

sufficiently small η , there exists a set $F \subset A$ with $\tilde{P}(F) > 0$ such that, for any $\xi > 0$, there exists T for which, on F ,

$$p_t > \eta, \quad \forall t > T \quad (13)$$

and

$$\tilde{E} \left[\sup_{s \geq t} \left\| \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] - \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| \middle| \mathcal{H}_{2t} \right] < \xi, \quad \forall t > T. \quad (14)$$

If Assumption 3 also holds, then for all $\psi > 0$,

$$\tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \psi \middle| \mathcal{H}_t \right\} \rightarrow 1, \quad (15)$$

uniformly on F .

Proof. Define the event $D_\eta = \{\omega \in A : \lim_{t \rightarrow \infty} p_t(\omega) > 2\eta\}$. Because the set A on which $\lim_{t \rightarrow \infty} p_t(\omega) > 0$ 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 $\|\tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}] - \hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}]\|$ tends \tilde{P} -almost surely to zero (by Lemma 2). Therefore, on D_η the random variable $Z_t = \sup_{s \geq t} \|\tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] - \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}]\|$ converges \tilde{P} -almost surely to zero and hence¹⁴

$$\tilde{E}[Z_t | \mathcal{H}_{2t}] \rightarrow 0 \quad \tilde{P} - \text{almost surely.} \quad (16)$$

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 $\tilde{E}[Z_t | \mathcal{H}_{2t}]$ is uniform. Hence, for any $\xi > 0$, there exists a time T such that the inequalities in (13) and (14) hold almost everywhere on F for all $t > T$.

Fix $\psi > 0$. Then, for all $\xi' > 0$, (14) holds for $\xi = \xi'\psi$, which implies that, uniformly on F ,

$$\tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \psi \middle| \mathcal{H}_{2t} \right\} \rightarrow 1.$$

But if Assumption 3 holds, then player 2's beliefs are measurable with respect to \mathcal{H}_t , so that (15) holds. \blacksquare

¹⁴The following implication is proved in Hart (1985, Lemma 4.24). Section A.3 reproduces the argument.

5.2 Proof of Theorem 1 (the “Easy” Case)

Since $\hat{\sigma}_1$ is never an equilibrium strategy in the long run, from Definition 2, 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 $BR(\hat{\sigma}_1)$.

Let $\gamma \equiv \min_{y,i,j} \rho_{ij}^y$, which is strictly positive from Assumption 1. Since $\hat{\sigma}_1$ is a strategy with a finite range, $\beta \equiv \min_{i,h_{1t}} \{\hat{\sigma}_1^i(h_{1t}) : \hat{\sigma}_1^i(h_{1t}) > 0\}$ is also strictly positive.

Suppose that there is a positive \tilde{P} -probability set of outcomes A on which $p_t \rightarrow 0$. Choose ξ, ζ such that $\zeta < \beta\gamma$ and $\xi < \min\{\psi, \beta - \zeta\gamma\}$, where ψ is the bound from Lemma 3. By (15), there is a \tilde{P} -positive measure set $F \subset A$ and $T \geq \bar{T}$ such that, on F and for any $t > T$,

$$\tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \xi \mid \mathcal{H}_t \right\} > 1 - \eta\zeta. \quad (17)$$

Moreover, Assumption 3 ensures that both $\hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}]$ and $\tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}]$ are in fact \mathcal{H}_t -measurable, and so (17) implies on F ,

$$\left\| \hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}] - \tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}] \right\| < \xi \quad \tilde{P} \text{ a.s.} \quad (18)$$

Let

$$g_t = \tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \xi \mid \mathcal{H}_{1t} \right\}.$$

Using the fact that $\{\mathcal{H}_{1t}\}_t$ is a finer filtration than $\{\mathcal{H}_t\}_t$ for the first equality, we have

$$\begin{aligned} & \tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \xi \mid \mathcal{H}_t \right\} \\ &= \tilde{E}[g_t | g_t \leq 1 - \eta, \mathcal{H}_t] \tilde{P}(g_t \leq 1 - \eta | \mathcal{H}_t) + \tilde{E}[g_t | g_t > 1 - \eta, \mathcal{H}_t] \tilde{P}(g_t > 1 - \eta | \mathcal{H}_t) \\ &\leq (1 - \eta) \tilde{P}(g_t \leq 1 - \eta | \mathcal{H}_t) + \tilde{P}(g_t > 1 - \eta | \mathcal{H}_t) \\ &= 1 - \eta + \eta \tilde{P}(g_t > 1 - \eta | \mathcal{H}_t), \end{aligned}$$

and so on F ,

$$1 - \eta\zeta < 1 - \eta + \eta \tilde{P}(g_t > 1 - \eta | \mathcal{H}_t),$$

i.e.,

$$\tilde{P}(g_t > 1 - \eta | \mathcal{H}_t) > 1 - \zeta,$$

or equivalently,

$$\tilde{P} \left(\tilde{P} \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \xi \mid \mathcal{H}_{1t} \right\} > 1 - \eta \mid \mathcal{H}_t \right) > 1 - \zeta, \quad (19)$$

i.e., 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. Because the commitment strategy is public, Lemma 3 ensures that player 2 plays a best response to the commitment strategy whenever believing that 1's strategy is within ξ of the commitment strategy. Hence, player 2 assigns a probability of at least $1 - \zeta$ to player 1 believing that player 2 is best responding to $\hat{\sigma}_1$ with at least probability $1 - \eta$.

We now argue that there is a period $t \geq T$ and an outcome in F such that $\hat{\sigma}_1$ is not optimal 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, \dots, y_{t+k} , $0 \leq k \leq K$, under which $\hat{\sigma}_1(h_t, y_t, \dots, y_{t+k})$ 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 $(h_t, y_t, \dots, y_{t+k})$ 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 $\ell \rightarrow \infty$, 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 (19), 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 (18).

5.3 Proof of Theorem 2 (the Harder Case)

If players 1 and 2 are not symmetrically informed about 2's beliefs, then player 1 needs to know player 2's private history h_{2s} in order to predict 2's

period- s beliefs and hence behavior. Unfortunately for player 1, she knows only her own private history h_{1s} .

We begin by identifying a circumstance under which the value of knowing player 2's private history becomes quite small. In general, telling player 1 the private history h_{2t} that player 2 observed through period $t < s$ will be of use in helping 1 predict 2's period- s behavior. However, the following Lemma shows that for a given t , h_{2t} becomes of arbitrarily small use in predicting player 2's actions in period s as $s \rightarrow \infty$.

The intuition is straightforward. Suppose first that h_{2t} is essential to predicting player 2's behavior in all periods $s > t$. Then, as time passes player 1 will figure out that h_{2t} actually occurred from her own observations. Hence, player 1 continues to receive information about this history from subsequent observations, reducing the value of having h_{2t} explicitly revealed. On the other hand if h_{2t} is of less and less use in predicting current behavior, then there is eventually no point in player 1 using it to predict player 2's behavior, again reducing the value of having h_{2t} revealed. In either case player 1's own period- s information swamps the value of learning h_{2t} , in the limit as s grows large.

Denote by $\beta(\mathcal{A}, \mathcal{B})$ the coarsest common refinement of the σ -algebras \mathcal{A} and \mathcal{B} . Thus, $\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})$ is the σ -algebra describing player 1's information at time s if she were to learn the private history of player 2 at time t . We also write $\beta(\mathcal{A}, B)$ for $\beta(\mathcal{A}, \{B, B^c\})$.

Lemma 5 *Suppose assumptions 1 and 2 hold. For any $t > 0$ and $\tau \geq 0$,*

$$\lim_{s \rightarrow \infty} \left\| \tilde{E}[\sigma_{2,s+\tau} | \beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \tilde{E}[\sigma_{2,s+\tau} | \mathcal{H}_{1s}] \right\| = 0, \quad \tilde{P}\text{-a.s. .}$$

Proof. We first provide the proof for the case of $\tau = 0$. Suppose that $K \subset J^t$ is a set of t -period player 2 action profiles $(j_0, j_1, \dots, j_{t-1})$. We also denote by K the event (i.e., subset of Ω) that the history of the first t -periods of player 2's action profiles are in K . By Bayes' rule and the finiteness of the action and signal spaces, we can write the conditional probability of the event K given the observation by player 1 of $h_{1,s+1} = (h_{1s}, y_s, i_s)$ as follows

$$\begin{aligned} \tilde{P}[K | h_{1,s+1}] &= \tilde{P}[K | h_{1s}, y_s, i_s] \\ &= \frac{\tilde{P}[K | h_{1s}] \tilde{P}[y_s, i_s | K, h_{1s}]}{\tilde{P}[y_s, i_s | h_{1s}]} \\ &= \frac{\tilde{P}[K | h_{1s}] \sum_j \rho_{i_s j}^{y_s} \tilde{E}[\sigma_2^j(h_{2s}) | h_{1s}, K]}{\sum_j \rho_{i_s j}^{y_s} \tilde{E}[\sigma_2^j(h_{2s}) | h_{1s}]}, \end{aligned}$$

where the last equality uses $\tilde{P}[i_s|K, h_{1s}] = \tilde{P}[i_s|h_{1s}]$.

Subtract $\tilde{P}[K|h_{1s}]$ from both sides to obtain

$$\tilde{P}[K|h_{1,s+1}] - \tilde{P}[K|h_{1s}] = \frac{\tilde{P}[K|h_{1s}] \sum_j \rho_{i_s j}^{y_s} \left(\tilde{E}[\sigma_2^j(h_{2s})|h_{1s}, K] - \tilde{E}[\sigma_2^j(h_{2s})|h_{1s}] \right)}{\sum_j \rho_{i_s j}^{y_s} \tilde{E}[\sigma_2^j(h_{2s})|h_{1s}]}.$$

The term $\sum_j \rho_{i_s j}^{y_s} \tilde{E}[\sigma_2^j(h_{2s})|h_{1s}]$ is player 1's conditional probability of observing the period- s signal y_s given she takes action i_s and hence is strictly positive and less than one by Assumption 1. Thus, we get,

$$\left| \tilde{P}[K|h_{1,s+1}] - \tilde{P}[K|h_{1s}] \right| \geq \tilde{P}[K|h_{1s}] \left| \sum_j \rho_{i_s j}^{y_s} \left(\tilde{E}[\sigma_2^j(h_{2s})|h_{1s}, K] - \tilde{E}[\sigma_2^j(h_{2s})|h_{1s}] \right) \right|.$$

The random variable $\tilde{P}[K|\mathcal{H}_{1s}]$ is a martingale with respect to the filtration $\{\mathcal{H}_{1s}\}$. Consequently it converges almost surely as $s \rightarrow \infty$ and hence the left side of this inequality converges almost surely to zero.¹⁵ The signals generated by player 2's actions satisfy Assumption 2, so an identical argument to that given at the end of Lemma 2 establishes that for \tilde{P} -almost all $\omega \in K$,

$$\lim_{s \rightarrow \infty} \tilde{P}[K|\mathcal{H}_{1s}] \left\| \tilde{E}[\sigma_{2s}|\beta(\mathcal{H}_{1s}, K)] - \tilde{E}[\sigma_{2s}|\mathcal{H}_{1s}] \right\| = 0.$$

The probability $\tilde{P}[K|\mathcal{H}_{1s}]$ is a martingale on the filtration $\{\mathcal{H}_{1s}\}_s$ with respect to \tilde{P} , and so \tilde{P} -almost surely converges to a nonnegative limit, $\tilde{P}[K|\mathcal{H}_{1\infty}]$. Moreover, $\tilde{P}[K|\mathcal{H}_{1\infty}](\omega) > 0$ for \tilde{P} -almost all $\omega \in K$. Thus, for \tilde{P} -almost all $\omega \in K$,

$$\lim_{s \rightarrow \infty} \left\| \tilde{E}[\sigma_{2s}|\beta(\mathcal{H}_{1s}, K)] - \tilde{E}[\sigma_{2s}|\mathcal{H}_{1s}] \right\| = 0.$$

Since this holds for all K ,

$$\lim_{s \rightarrow \infty} \left\| \tilde{E}[\sigma_{2s}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \tilde{E}[\sigma_{2s}|\mathcal{H}_{1s}] \right\| = 0, \quad \tilde{P}\text{-a.s.},$$

giving the result for $\tau = 0$.

The proof for $\tau \geq 1$ follows by induction. In particular, we have

$$\begin{aligned} \Pr[K|h_{1,s+\tau+1}] &= \Pr[K|h_{1s}, y_s, i_s, \dots, y_{s+\tau}, i_{s+\tau}] \\ &= \frac{\Pr[K|h_{1s}] \Pr[y_s, i_s, \dots, y_{s+\tau}, i_{s+\tau}|K, h_{1s}]}{\Pr[y_s, i_s, \dots, y_{s+\tau}, i_{s+\tau}|h_{1s}]} \\ &= \frac{\Pr[K|h_{1s}] \prod_{z=s}^{s+\tau} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}, K]}{\prod_{z=s}^{s+\tau} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}]}, \end{aligned}$$

¹⁵See footnote 12.

where $h_{2,z+1} = (h_{2z}, y_z, i_z)$. Hence,

$$\begin{aligned} & |\Pr[K|h_{1,s+\tau+1}] - \Pr[K|h_{1s}]| \\ & \geq \Pr[K|h_{1s}] \left| \prod_{z=s}^{s+\tau} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}, K] - \prod_{z=s}^{s+\tau} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}] \right|. \end{aligned}$$

The left side of this inequality converges to zero \tilde{P} -almost surely, and hence so does the right side. Moreover, the right side is larger than

$$\begin{aligned} & \Pr[K|h_{1s}] \left| \prod_{z=s}^{s+\tau-1} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}] \right| \\ & \quad \times \left| \sum_j \rho_{i_{s+\tau} j}^{y_{s+\tau}} \tilde{E}[\sigma_2^j(h_{2,s+\tau})|h_{1s}, K] - \sum_j \rho_{i_{s+\tau} j}^{y_{s+\tau}} \tilde{E}[\sigma_2^j(h_{2,s+\tau})|h_{1s}] \right| \\ & - \Pr[K|h_{1s}] \left| \prod_{z=s}^{s+\tau-1} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}, K] - \prod_{z=s}^{s+\tau-1} \sum_j \rho_{i_z j}^{y_z} \tilde{E}[\sigma_2^j(h_{2z})|h_{1s}] \right| \\ & \quad \times \left| \sum_j \rho_{i_{s+\tau} j}^{y_{s+\tau}} \tilde{E}[\sigma_2^j(h_{2,s+\tau})|h_{1s}, K] \right|. \end{aligned}$$

From the induction hypothesis that $\|\tilde{E}[\sigma_{2z}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \tilde{E}[\sigma_{2z}|\mathcal{H}_{1s}]\|$ converges to zero \tilde{P} -almost surely for every $z \in \{s, \dots, s + \tau - 1\}$, the negative term also converges to zero \tilde{P} -almost surely. But then the first term also converges to zero, and, as above, the result holds for $z = s + \tau$. \blacksquare

If the \tilde{P} limit of player 2's posterior belief p_t is not zero, then (from Lemma 4) there must be a positive probability set of states of the world and a time T such that at time T , player 2 thinks he will play a best response to the commitment type in all future periods (with arbitrarily high probability). At time T , player 1 may not realize player 2 has such a belief. However, there is a positive measure set of states where an observer at time s , given the finer information partition $\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})$, attaches high probability to player 2 playing a best response to the commitment type forever. However, by Lemma 5 we can then deduce that player 1 will also attach high probability to player 2 playing a best response to the commitment type for s large:

Lemma 6 *Suppose Assumptions 1 and 2 hold and $\hat{\sigma}_1$ is realization equivalent to a public strategy. Suppose, moreover, that there exists a \tilde{P} -positive measure set of states of the world for which $\lim_{t \rightarrow \infty} p_t(\omega) > 0$. Then, for any $\nu > 0$ and any integer $\tau > 0$ there exists a \tilde{P} -positive measure set of states F^\dagger and a time $T(\nu, \tau)$ for which, for all $s > T(\nu, \tau)$, for all $\omega \in F^\dagger$ and any $s' \in \{s, s+1, \dots, s+\tau\}$,*

$$\min_{\hat{\sigma}_2 \in BR(\hat{\sigma}_1)} \left\| \tilde{E}[\sigma_{2s'} | \mathcal{H}_{1s}] - \hat{\sigma}_{2s'} \right\| < \nu, \quad \tilde{P}\text{-a.s.}$$

Proof. The first step of this proof establishes the existence of a positive probability event ($\tilde{P}(F) > 0$) such that, for all t large, player 2 attaches high probability to always playing a best response to the commitment type in all future periods almost surely on F .

Define the event $D_\eta = \{\omega : \lim_{t \rightarrow \infty} p_t(\omega) > 2\eta\}$. From Lemma 4, for sufficiently small η there exists $F \subset D_\eta$ such that $\tilde{P}(F) = \mu$ for some $\mu > 0$ and such that, for any $\xi < \{\mu^2, \nu^2/9\}$, there exists T for which, on F and $\forall t > T$,

$$p_t > \eta, \quad (20)$$

and

$$\tilde{E}[\sup_{s \geq t} \|\tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] - \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}]\| | \mathcal{H}_{2t}] < \xi\psi,$$

where $\psi > 0$ is given by Lemma 3. Let $Z_t = \sup_{s \geq t} \left\| \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] - \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] \right\|$. As $\tilde{E}[Z_t | \mathcal{H}_{2t}] < \xi\psi$ for all $t > T$ on F and $Z_t \geq 0$, $\tilde{P}(\{Z_t > \psi\} | \mathcal{H}_{2t}) < \xi$ for all $t > T$ on F . This and Lemma 3 imply that, almost everywhere on F ,

$$\tilde{P}(\{\sigma_{2s} \in BR(\hat{\sigma}_{1s}), \forall s > t\} | \mathcal{H}_{2t}) > 1 - \xi, \quad \forall t > T. \quad (21)$$

This last argument uses the condition that $\hat{\sigma}_1$ is realization equivalent to a public strategy, ensuring that $\hat{\sigma}_{1s}$ is measurable with respect to player 2's filtration.

The second step in the proof shows that there is a positive probability set of states F^* where (21) holds and where player 1 must also believe that player 2 is playing a best response to the commitment type with high probability. Now we define two types of event:

$$G_t \equiv \{\omega : \sigma_{2s} \in BR(\hat{\sigma}_{1s}), \forall s \geq t\}$$

and

$$K_t \equiv \left\{ \omega : \tilde{P}(G_t | \mathcal{H}_{2t}) > 1 - \xi \right\}.$$

Note that $K_t \in \mathcal{H}_{2t}$ and $F \subset \cap_{t>T} K_t$. For a given t define the random variable g_s to be the \tilde{P} -probability of the event G_t conditional on the private history h_{1s} of player 1 and the private history h_{2t} of player 2, that is¹⁶

$$g_s \equiv \tilde{E} [\mathbf{1}_{G_t} \mid \beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})],$$

where $\mathbf{1}_{G_t}$ is the indicator function for the set G_t . However, $\tilde{P}(G_t | \mathcal{H}_{2t}) = \tilde{E}[g_s | \mathcal{H}_{2t}]$ and

$$\begin{aligned} \tilde{E}[g_s | \mathcal{H}_{2t}] &= \tilde{E}[g_s | g_s \leq 1 - \sqrt{\xi}, \mathcal{H}_{2t}] \tilde{P}(g_s \leq 1 - \sqrt{\xi} | \mathcal{H}_{2t}) \\ &\quad + \tilde{E}[g_s | g_s > 1 - \sqrt{\xi}, \mathcal{H}_{2t}] \tilde{P}(g_s > 1 - \sqrt{\xi} | \mathcal{H}_{2t}) \\ &\leq (1 - \sqrt{\xi}) \tilde{P}(g_s \leq 1 - \sqrt{\xi} | \mathcal{H}_{2t}) + \tilde{P}(g_s > 1 - \sqrt{\xi} | \mathcal{H}_{2t}) \\ &= (1 - \sqrt{\xi}) + \sqrt{\xi} \tilde{P}(g_s > 1 - \sqrt{\xi} | \mathcal{H}_{2t}). \end{aligned}$$

For every $\omega \in K_t \in \mathcal{H}_{2t}$ it is the case that $\tilde{P}(G_t | \mathcal{H}_{2t}) > 1 - \xi$. Thus

$$\tilde{P}(g_s > 1 - \sqrt{\xi} | \mathcal{H}_{2t}) > \frac{(1 - \xi) - (1 - \sqrt{\xi})}{\sqrt{\xi}} = 1 - \sqrt{\xi}, \quad \forall \omega \in K_t. \quad (22)$$

As $F \subset K_t$, (22) holds for all $\omega \in F$. The random variable g_s is a bounded martingale with respect to the filtration $\{\mathcal{H}_{2s}\}$ and so converges almost surely to a limiting random variable, which we denote by g_∞ . There is a \tilde{P} -positive measure set $F^* \subset F$ and a time T_t such that for $s > T_t$, $g_s(\omega) > 1 - 2\sqrt{\xi}$ on F^* .¹⁷ To summarize — for almost every state on F^* and all $t > T$ and all $s > T_t > T$:

$$\tilde{E} [\mathbf{1}_{G_t} \mid \mathcal{H}_{2t}] > 1 - \xi, \quad \tilde{P}\text{-a.s.}$$

and

$$\tilde{E} [\mathbf{1}_{G_t} \mid \beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] > 1 - 2\sqrt{\xi}, \quad \tilde{P}\text{-a.s.} \quad (23)$$

Finally, by Lemma 5 and an application of Egorov's Theorem, for any $\xi > 0$, $\zeta > 0$ and τ there exists $F^\dagger \subset F^*$ satisfying $\tilde{P}(F^*) - \tilde{P}(F^\dagger) < \zeta$ and T_τ such that for all $s > T_\tau$ and $s' = s, s + 1, \dots, s + \tau$,

$$\left\| \tilde{E}[\sigma_{2s'} | \beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \tilde{E}[\sigma_{2s'} | \mathcal{H}_{1s}] \right\| < \sqrt{\xi}, \quad \tilde{P}\text{-a.s.} \quad (24)$$

¹⁶Recall that we use $\beta(\mathcal{A}, \mathcal{B})$ to denote the σ -algebra that is the coarsest common refinement of the σ -algebras \mathcal{A} and \mathcal{B} .

¹⁷Proof: As $F \subset K_t$ and $\mu = \tilde{P}(F)$ we have that $\tilde{P}(F | K_t) = \mu / \tilde{P}(K_t)$. However, if $g_\infty < 1 - \sqrt{\xi}$ for almost every $\omega \in F$, then $\tilde{P}(F | K_t) \leq \tilde{P}(g_\infty < 1 - \sqrt{\xi} | K_t)$. By (22) however, this implies that $\tilde{P}(F | K_t) \leq \sqrt{\xi}$. Combining $\tilde{P}(F | K_t) = \mu / \tilde{P}(K_t)$ and $\tilde{P}(F | K_t) \leq \sqrt{\xi}$, we get $\mu / \sqrt{\xi} \leq \tilde{P}(K_t)$, contradicting $\xi < \mu^2$. Hence, there exists a positive probability subset of F and a T_t such that on which $g_s \geq 1 - 2\sqrt{\xi}$ \tilde{P} -a.s. for all $s > T_t$.

Now $\tilde{E}[\sigma_{2s'}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] = g_s \hat{\sigma}_{2s'} + (1 - g_s) \tilde{E}[(1 - \mathbf{1}_{G_t})\sigma_{2s'}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})]$ for some $\hat{\sigma}_2 \in BR(\hat{\sigma}_1)$. Hence

$$\left\| \tilde{E}[\sigma_{2s'}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \hat{\sigma}_{2s'} \right\| \leq |1 - g_s|, \quad \tilde{P}\text{-a.s.} \quad (25)$$

Now, combine (24) and (25) with (23). This gives the result that there is a T and a $T' = \max\{T_t, T_\tau\}$ and F^\dagger such that for all $t > T$, all $s > T_t > T$ and $s' = s, s + 1, \dots, s + \tau$:

$$\min_{\hat{\sigma}_2 \in BR(\hat{\sigma}_1)} \left\| \tilde{E}[\sigma_{2s'}|\mathcal{H}_{1s}] - \hat{\sigma}_{2s'} \right\| < 3\sqrt{\xi}, \quad \tilde{P}\text{-a.s. .}$$

Given $3\sqrt{\xi} < \nu$ we have proved the lemma. \blacksquare

We can now prove Theorem 2. Intuitively, suppose the theorem is false and hence there is a positive measure set of states on which the limiting posterior p_∞ is positive. Since $\hat{\sigma}_1$ is a pure strategy, it is realization equivalent to a pure public strategy. By Lemma 6 there is a set of states where, for all s sufficiently large, the normal type attaches high probability to player 2 best responding to the commitment type for the next τ periods. The normal type's best response is not the commitment strategy, which is never an equilibrium in the long run. At an equilibrium, therefore, the normal type is *not* playing the commitment strategy on this set of states. From Lemma 2, however, we know that player 2 believes the average strategy played by the normal type is close to the strategy played by the commitment type whenever the limiting posterior p_∞ is positive. Moreover, if the commitment type is playing a pure strategy, the expected strategy of the normal type can only converge to the commitment type's strategy if the probability (conditional upon player 2's information) that the normal type is *not* playing the commitment action is vanishing small. This will give a contradiction.

More formally, suppose there exists an equilibrium where $p_\infty > 0$ on a set of strictly positive \tilde{P} -measure. Choose $M > \max_{\ell \in \{1,2\}, i \in I, j \in J} |\pi_\ell(i, j)|$. Choose τ sufficiently large that $\delta^\tau 32M < \varepsilon$, where ε is given by Definition 2. Also, choose ν sufficiently small for $14\varepsilon > (\tau + 1)\nu(32M + 15\varepsilon)$. By Lemma 6, there exists F^\dagger and a time $T(\nu, \tau)$ such for all $t > T(\nu, \tau)$ the normal type believes that player 2 is playing within ν of a best response to the commitment type over the next τ periods.

By our choice of τ any change in player 1's strategy after $t + \tau$ periods will change her expected discounted payoff at time t by at most $\varepsilon/16$. However, if player 2 always plays a best response to $\hat{\sigma}_1$, which is never an equilibrium in the long run, then there exists a deviation from $\hat{\sigma}_1$ that yields player 1

an increased discounted expected payoff of ε . This deviation must increase player 1's expected payoff by at least $\frac{15}{16}\varepsilon$ over the first τ periods. On F^\dagger for $t > T(\nu, \tau)$, player 1 attaches probability at least $1 - (\tau + 1)\nu$ to player 2 playing a best response to $\hat{\sigma}_1(h_{1s'})$ in periods $s' = t, t+1, \dots, t+\tau$. (Since the commitment type's strategy is public, the set of player 2's best responses to $\hat{\sigma}_1(h_{1s'})$ is public, and any belief that player 2 is playing some best reply to $\hat{\sigma}_1(h_{1s'})$ is equivalent to a point belief that player 2 is playing a particular, possibly randomized, best reply.) Thus her expected gain from the deviation is at least $(1 - (\tau + 1)\nu)\frac{15}{16}\varepsilon + (\tau + 1)\nu(-2M)$, which exceeds $\frac{1}{16}\varepsilon$ (by our choice of ν), which is the largest payoff gain player 1 can achieve by adhering to the commitment strategy for τ periods and playing optimally thereafter. Hence, on F^\dagger for all $t > T(\nu, \tau)$, the continuation strategy $\hat{\sigma}_1$ is not a best reply to the expected behavior of player 2.

Let F_t^s denote the subset of F^\dagger where the normal type's strategy first puts zero probability on the commitment action in period $s \geq t$, i.e., if $i_z(\omega) \in I$ denotes the action specified by the pure strategy $\hat{\sigma}_1$ in period z in state ω ($\hat{\sigma}_{1z}^{i_z(\omega)}(\omega) = 1$), then $F_t^s \equiv \{\omega \in F^\dagger : \tilde{\sigma}_{1s}^{i_s(\omega)}(\omega) = 0, \tilde{\sigma}_{1z}^{i_z(\omega)}(\omega) > 0 \forall z = t, t+1, \dots, s-1\}$. Then, for infinitely many $t > \max\{T(\nu, \tau), \bar{T}\}$, $\tilde{P}(F^\dagger \setminus \bigcup_{s=t}^{t+\tau} F_t^s) = 0$, where \bar{T} is the bound from Definition 2. (The argument is identical to that in the penultimate paragraph of the proof of Theorem 1.)

The remainder of the proof argues on the subsequence of t just identified. If we choose s to maximize $\tilde{P}(F_t^s)$, then $\tilde{P}(F_t^s) \geq \tilde{P}(F^\dagger)/(\tau + 1)$. Moreover, there is a subset $F^\ddagger \subset F^\dagger$ with $\tilde{P}(F^\ddagger) > 0$, and an increasing sequence of dates, $\{s_m\}$, such that for all s_m on F^\ddagger ,

$$\|\hat{\sigma}_{1s_m} - \tilde{\sigma}_{1s_m}\| = 1.$$

Since $p_t > \eta$ on F^\dagger , by (20), on F^\ddagger ,

$$p_{s_m} \|\hat{\sigma}_{1s_m} - \tilde{\sigma}_{1s_m}\| \geq \eta.$$

As the random variable $p_{s_m} \|\hat{\sigma}_{1s_m} - \tilde{\sigma}_{1s_m}\|$ is non-negative,

$$\tilde{E}[p_{s_m} \|\hat{\sigma}_{1s_m} - \tilde{\sigma}_{1s_m}\|] \geq \eta \tilde{P}(F^\ddagger) > 0. \quad (26)$$

From Corollary 1, we have

$$\lim_{t \rightarrow \infty} p_t \|\hat{\sigma}_{1t} - \tilde{E}[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}]\| = 0, \quad \tilde{P}\text{-a.s.}$$

This uses the fact that $\hat{\sigma}_1$ is (realization equivalent to) a public strategy and so $\hat{\sigma}_{1t} = \hat{E}[\hat{\sigma}_{1t} | \mathcal{H}_{2t}]$. Let $i(h_{2t})$ be the pure action played by the commitment

strategy after the history h_{2t} , that is, $\hat{\sigma}_{1t}^{i(h_{2t})}(\omega) = 1$. (This is observable by player 2 at time t if $\hat{\sigma}_1$ is pure.) Then, as $\|\cdot\|$ is the supremum norm and both vectors are contained in Δ^I ,

$$\begin{aligned}\|\hat{\sigma}_{1t} - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_{2t}]\| &= 1 - \tilde{E}[\tilde{\sigma}_{1t}^{i(\cdot)}|\mathcal{H}_{2t}] \\ &= \tilde{E}[1 - \tilde{\sigma}_{1t}^{i(\cdot)}|\mathcal{H}_{2t}] \\ &= \tilde{E}[\|\hat{\sigma}_{1t} - \tilde{\sigma}_{1t}\||\mathcal{H}_{2t}].\end{aligned}$$

Since p_t is measurable with respect to the filtration $\{\mathcal{H}_{2t}\}_t$, we have

$$\lim_{t \rightarrow \infty} \tilde{E}[p_t \|\hat{\sigma}_{1t} - \tilde{\sigma}_{1t}\||\mathcal{H}_{2t}] = 0, \quad \tilde{P}\text{-a.s.}$$

Now, taking an unconditional expectation of the conditional expectation, we get

$$\lim_{t \rightarrow \infty} \tilde{E}[p_t \|\hat{\sigma}_{1t} - \tilde{\sigma}_{1t}\|] = 0,$$

contradicting (26).

6 Imperfect Private Monitoring

In this section, we briefly consider the case of imperfect private monitoring. At the end of the period, player 1 observes a private signal θ (drawn from a finite signal space Θ) of the action profile chosen in that period. Similarly, player 2 observes a private signal ζ (drawn from a finite signal space Z) of the action profile chosen in that period. Given the underlying action profile (i, j) , we let ρ_{ij} be a probability distribution over $\Theta \times Z$, so that $\rho_{ij}^{\theta\zeta}$ is the probability that the signal profile (θ, ζ) is observed. The marginal distributions are given by $\rho_{ij}^\theta = \sum_\zeta \rho_{ij}^{\theta\zeta}$ and $\rho_{ij}^\zeta = \sum_\theta \rho_{ij}^{\theta\zeta}$. If $\Theta = Z$ and $\sum_{\theta \in \Theta} \rho_{ij}^{\theta\theta} = 1$ for all i, j , the monitoring is *public*. A particular signal realization $\theta'\zeta' \in \Theta \times Z$ is *public* if for all i and j , $\rho_{ij}^{\theta'\zeta'} > 0$, and $\rho_{ij}^{\theta'\zeta} = 0$ and $\rho_{ij}^{\theta\zeta'} = 0$ for all $\zeta \neq \zeta'$ and $\theta \neq \theta'$. Histories for the players are defined in the obvious way. The full-support assumption is:

Assumption 5 $\rho_{ij}^\theta, \rho_{ij}^\zeta > 0$ for all $\theta \in \Theta, \zeta \in Z$, and $(i, j) \in I \times J$.

Note that we do *not* assume that $\rho_{ij}^{\theta\zeta} > 0$ for all $(i, j) \in I \times J$ and $(\theta, \zeta) \in \Theta \times Z$ (which would rule out public monitoring). Rather the full-support assumption is that each signal is observed with positive probability under every action profile.

Even when monitoring is *truly private*, in the sense that $\rho_{ij}^{\theta\zeta} > 0$ for all $(i, j) \in I \times J$ and $(\theta, \zeta) \in \Theta \times Z$, reputations can have very powerful short-run effects. While Fudenberg and Levine (1992) explicitly assume the game has public monitoring, under the following identification assumption, the analysis of Fudenberg and Levine (1992) covers imperfect private monitoring, implying the following Theorem 5.

Assumption 6 For all $j \in J$, there are I linearly independent columns in the matrix $(\rho_{ij}^{\zeta})_{\zeta \in Z, i \in I}$.

Theorem 5 Suppose the game has imperfect private monitoring satisfying Assumptions 5 and 6. Suppose the commitment type is a simple action type that plays the pure action i^* in every period. For all priors $p_0 > 0$ and all $\varepsilon > 0$, there exists $\bar{\delta} \in (0, 1)$ such that for all $\delta \in n(\bar{\delta}, 1)$, player 1's expected average discounted payoff in any Nash equilibrium is at least

$$\min_{j \in BR(i^*)} \pi_1(i^*, j) - \varepsilon,$$

where

$$BR(i) = \operatorname{argmax}_{j \in J} \pi_2(i, j).$$

Assumption 6 covers the setting where the signal of player 1's action is only observed by player 2, and player 1 observes no signal.¹⁸ When player 1 also receives signals, we will need an assumption analogous to Assumption 4:

Assumption 7 For all $i \in I$, there are J linearly independent columns in the matrix $(\rho_{ij}^{\theta})_{\theta \in \Theta, j \in J}$.

The result (the proof is essentially that of Theorem 2) in the setting with imperfect private monitoring is then:

Theorem 6 Suppose the imperfect private monitoring distribution ρ satisfies Assumptions 5, 6, and 7. Suppose $\hat{\sigma}_1$ is a pure public strategy that is never an equilibrium strategy in the long run. In any Nash equilibrium, $p_t \rightarrow 0$ \tilde{P} -almost surely.

¹⁸In this case, when there is complete information, the one-period memory strategy profiles that we describe as equilibria in Section 2 are also equilibria of the game with private monitoring. We thank Juuso Valimaki for showing us how to construct such equilibria.

If the monitoring is truly private, then there is no public information. In that case, a pure public strategy is simple, i.e., the same pure action taken in every period.¹⁹ This is, of course, the typical type studied in applications.

Finally, it is straightforward to verify that Theorem 3 in Section 4.3 extends to the case of imperfect private monitoring.

7 Many Types

It is straightforward to extend the preceding analysis to case of many types.

Let \mathcal{T} be the set of possible commitment types. The commitment type c plays the fixed, repeated-game strategy $\hat{\sigma}_1^c$. We assume \mathcal{T} is either finite or countably infinite. 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 $p_0^n = 1 - \sum_{c \in \mathcal{T}} p_0^c$ 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 .

Definition 5 *A set of commitment types \mathcal{T} is separated if for all $c, c' \in \mathcal{T}$,*

$$\frac{\hat{P}^{c'}(h_{2t})}{\hat{P}^c(h_{2t})} \rightarrow 0 \quad \hat{P}^c\text{-a.s.}$$

In other words, a set of commitment types is separated if player 2 can always learn which commitment type he faces, if he knows he is facing one of them.

We need the following Lemma.

Lemma 7 *For all c and c' in a separated set of commitment types, in any Nash equilibrium, \tilde{P} -almost surely,*

$$p_t^c p_t^{c'} \rightarrow 0.$$

Proof. We argue to a contradiction. Suppose there exists a set $F \subset \Omega$ such that $\tilde{P}(F) > 0$, and on F , for c and c' in a separated set of commitment types,

$$\lim_{t \rightarrow \infty} p_t^c > 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} p_t^{c'} > 0. \quad (27)$$

¹⁹Notice that if the monitoring is not public, then a pure strategy for player 1 need not be realization equivalent to a public strategy, prompting the requirement in Theorem 6 that player 1’s strategy be pure public.

We first argue that $\hat{P}^c(F) = 0$. If $\hat{P}^c(F) > 0$, since the commitment types are separated, on a full measure subset of F ,

$$\frac{\hat{P}^{c'}(h_{2t})}{\hat{P}^c(h_{2t})} \rightarrow 0.$$

But

$$p_t^{c'} = \frac{\hat{P}^{c'}(h_{2t})}{\hat{P}^c(h_{2t})} p_0^{c'} p_t^c,$$

and so on this full measure subset, $p_t^{c'} \rightarrow 0$, contradicting (27). Thus, $\hat{P}^c(F) = 0$.

However, since the unconditional probability of F is strictly positive,

$$0 < E\{p_t^c|F\} = \frac{\hat{P}^c(F) p_0^c}{P(F)}.$$

But $\hat{P}^c(F) = 0$ implies $E\{p_t^c|F\} = 0$, a contradiction. ■

We then have:

Theorem 7 *Suppose ρ satisfies Assumptions 1, 2, and 3. Suppose \mathcal{T} is a separated set of commitment types and the support of the prior p_0 is $\mathcal{T} \cup \{n\}$. Let \mathcal{T}^* be the set of commitment types $c \in \mathcal{T}$ for which $\hat{\sigma}_1^c$ is a public strategy with finite range that is never an equilibrium strategy in the long run. Then in any Nash equilibrium, $p_t^c \rightarrow 0$ for all $c \in \mathcal{T}^*$ \tilde{P} -almost surely.*

Theorem 8 *Suppose ρ satisfies Assumptions 1, 2, and 4. Suppose \mathcal{T} is a separated set of commitment types and the support of the prior p_0 is $\mathcal{T} \cup \{n\}$. Let \mathcal{T}^* be the set of commitment types $c \in \mathcal{T}$ for which $\hat{\sigma}_1^c$ is a pure strategy that is never an equilibrium strategy in the long run. Then in any Nash equilibrium, $p_t^c \rightarrow 0$ for all $c \in \mathcal{T}^*$ \tilde{P} -almost surely.*

The proofs of these two results are almost identical to the proofs of Theorems 1 and 2, with the following change. Fix some type $c' \in \mathcal{T}^*$. In the proofs, 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 each proof, when the normal type has a profitable deviation that contradicts player 2's beliefs. We now apply Lemma 7. 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 contradictions.

8 Two Long-Run Players

Lemma 3 is the only place where the assumption that player 2 is short-lived makes an appearance. When player 2 is short-lived, player 2 is 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)).

In order to extend our results to long-lived players we need an analog of Lemma 3. Again, for expositional clarity, we restrict attention to the case of a single commitment type.

Lemma 8 *Suppose $\hat{\sigma}_1$ is a public strategy implementable by a finite automaton, denoted (W, d, φ, w_0) , and $BR(\hat{\sigma}_1; w')$ is the set of best replies for player 2 to the public strategy implemented by the finite automaton (W, d, φ, w') , i.e., the initial state is $w' \in W$. For any history h_{2t} , let $w(h_{2t}) \in W$ be the state reached from w_0 under the public history consistent with h_{2t} . Let $(\tilde{\sigma}_1, \sigma_2)$ be equilibrium strategies in the incomplete-information game where player 2 is long-lived with discount factor $\delta_2 \in [0, 1)$. If σ_2 is a pure strategy, then for all $T > 0$ there exists $\psi > 0$ such that if player 2 observes a history h_{2t} so that*

$$P \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \psi \mid h_{2t} \right\} > 1 - \psi, \quad (28)$$

then for some $\sigma'_2 \in BR(\hat{\sigma}_1; w(h_{2t}))$, the continuation strategy of σ_2 after the history h_{2t} 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 4 shows that (28) holds (note that (14) in Lemma 4 uses \tilde{P} rather than P to evaluate the probability of the event of interest).

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 for such if player 2 observes a history h_{2t}

so that $w = w(h_{2t})$ and

$$P \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| < \psi_w \middle| h_{2t} \right\} > 1 - \psi_w, \quad (29)$$

then for some $\sigma'_2 \in BR(\hat{\sigma}_1; w)$, the continuation strategy of σ_2 after the history h_{2t} agrees with σ'_2 for the next T periods.

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

$$\left\| \hat{\sigma}_1(h_{2s}) - \tilde{E}[\tilde{\sigma}_{1s} | h_{2s}] \right\| < \varepsilon_w, \quad (30)$$

then for some $\sigma'_2 \in BR(\hat{\sigma}_1; w(h'_{2t}))$, the continuation strategy of σ_2 after the history h'_{2t} agrees with σ'_2 for the next T periods.

Recall that $\gamma \equiv \min_{y, ij} \rho_{ij}^y$ and set $\psi_w = \frac{1}{2} \min \left\{ \varepsilon_w, \gamma^{T'} \right\}$. Suppose (29) holds with this ψ_w . We claim that (30) holds for $s = t, \dots, t + T'$ and for all h_{2s} with initial segment h'_{2t} . Suppose not. Since player 2 is following a pure strategy, the probability of the continuation history h_{2s} , conditional on the history h'_{2t} , is at least $\gamma^{T'}$. Thus,

$$P \left\{ \sup_{s \geq t} \left\| \hat{E}[\hat{\sigma}_{1s} | \mathcal{H}_{2s}] - \tilde{E}[\tilde{\sigma}_{1s} | \mathcal{H}_{2s}] \right\| \geq \psi_w \middle| h_{2t} \right\} \geq \gamma^{T'},$$

contradicting (29), since $\gamma^{T'} > \psi_w$. ■

With this result in hand, the proofs of Theorems 1 and 2 go through as before, extending our result to two long-lived players, provided the commitment type is a public strategy implementable by a finite automaton and player 2 plays a pure strategy in the Nash equilibrium.

A Appendix

A.1 Proof of Theorem 3 (Section 4.3)

Proof. Since $p_t \rightarrow 0$ \tilde{P} -almost surely, we have $p_t \rightarrow 1$ \hat{P} -almost surely. For any $\varepsilon, \nu > 0$ there exists a T such that for all $t > T$, $\tilde{P}(p_t > \varepsilon) + \hat{P}(p_t <$

$1 - \varepsilon) < \nu$. Hence, for $t' > T$,

$$\begin{aligned} 0 &\leq \int_{S \times [0,1]} [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2(s_2, p_t))] d(p_0 \hat{\rho}_t + (1 - p_0) \tilde{\rho}_t) \\ &\leq (1 - p_0) \int_{S \times [0, \varepsilon]} [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2(s_2, p_t))] d\tilde{\rho}_t \\ &\quad + p_0 \int_{S \times [1 - \varepsilon, 1]} [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2(s_2, p_t))] d\hat{\rho}_t + 2M\nu, \end{aligned}$$

where M is an upper bound on the magnitude of the stage-game payoffs and the first inequality holds because we have a subjective correlated equilibrium. As ξ_2 is measurable with respect to p_t , we can ensure that the final integral in the preceding expression is zero by setting $\xi_2(s_2, p_t) = s_2$ for $p_t > \varepsilon$, and hence, for any $\varepsilon, \nu > 0$ and for all ξ_2 ,

$$\int_{S \times [0, \varepsilon]} [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \zeta_2(s_2, p_t))] d\tilde{\rho}_t \geq -\frac{2M\nu}{1 - p_0}, \quad \forall t' > T. \quad (\text{A.1})$$

Again, because $\tilde{P}(p_t > \varepsilon) < \nu$, (A.1) implies

$$\int_{S \times [0,1]} [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2(s_2, p_t))] d\tilde{\rho}_t \geq -\frac{2M\nu}{1 - p_0} - 2M\nu, \quad \forall t' > T.$$

Integrating out p_t implies that, for all $\xi_2' : S_2 \rightarrow S_2$,

$$\int_S [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2'(s_2))] d\tilde{\mu}_t \geq -\frac{2M\nu}{1 - p_0} - 2M\nu, \quad \forall t' > T. \quad (\text{A.2})$$

Consider now a convergent subsequence, denoted $\tilde{\mu}_{t_k}$ with limit $\tilde{\mu}_\infty$, and suppose $\tilde{\mu}_\infty$ is not a correlated equilibrium. Since (8) holds for all t' , it also holds in the limit, and so for some t' and some $\xi_2'' : S_2 \rightarrow S_2$, there exists $\kappa > 0$ so that

$$\int_S [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2''(s_2))] d\tilde{\mu}_\infty < -\kappa < 0.$$

But then for t_k sufficiently large,

$$\int_S [u_2^{t'}(s_1, s_2) - u_2^{t'}(s_1, \xi_2''(s_2))] d\tilde{\mu}_{t_k} < \frac{-\kappa}{2} < 0,$$

contradicting (A.2) for ν sufficiently small. ■

A.2 Proof of Theorem 4 (Section 4.4)

We begin by focusing on games that are “close” to the complete-information game.

Lemma A *Let (i^*, j^*) be a strict Nash equilibrium of the one-shot game. 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 $\varepsilon' = \frac{1}{2} [\pi_1(i^*, j^*) - \max_{i \neq i^*} \pi_1(i, j^*)] > 0$. Since the Nash equilibrium correspondence is upper hemicontinuous, 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.

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^* . Now, for any T , there exists $\hat{\eta} > 0$ so that if $p_0 < \hat{\eta}$, $\max p_t < \min\{\eta', \eta''\}$ for all $t \leq T$. 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. ■

While 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 *Let (i^*, j^*) be a strict Nash equilibrium of the one-shot game. 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.

By Theorem 1, $p_t \rightarrow 0$ \tilde{P} -almost surely, and so $\sup_{t' \geq t} p_{t'} \rightarrow 0$ \tilde{P} -almost surely. By Egorov's Theorem, there exists a t^* such that $\tilde{P} \{ \sup_{t' \geq t^*} p_{t'} > \eta'' \} < \varepsilon$. But then for some history for player 2, h_{2t^*} , $\tilde{P} \{ \sup_{t' \geq t^*} p_{t'} > \eta'' | h_{2t^*} \} < \varepsilon$. By the construction of the equilibrium, the continuation play after such a history (which necessarily leads to a belief, $p_{t^*}(h_{2t^*})$, for player 2 satisfying $p_{t^*}(h_{2t^*}) \leq \eta''$) is identical to that in the incomplete-information game with initial prior $p_0 = p_{t^*}(h_{2t^*})$. Thus, for the incomplete-information game with prior $p_0 = p_{t^*}(h_{2t^*})$, the probability that the posterior after any history exceeds η'' is no more than ε .

The proof is completed by setting $\eta = p_{t^*}(h_{2t^*})$, since for all t and all h_{2t} , we have $p_t(h_{2t}; p'_0) < p_t(h_{2t}; p_0)$ for any prior $p_0 < p_{t^*}(h_{2t^*})$. ■

We can then prove Theorem 4:

Proof. We prove this by first constructing an equilibrium of an artificial game, and then arguing 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 $\{(i_s, j_s, g)_{s=0}^{t-1}, (i_t, j_t, e, p_t)\}$, 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.

By construction, player 2 always finds it strictly optimal to announce his posterior.²⁰ Moreover, again by construction, player 2 ends the game if and only if his posterior is less than η .

Now consider an equilibrium (σ_1^*, σ_2^*) of the artificial game.²¹ Then let play in the original game be 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 $1 - \varepsilon$. ■

A.3 Verification of (16) (Section 5.1.3)

Lemma C *Suppose $\{X_m\}$ is a bounded sequence of random variables, and $X_m \rightarrow 0$ almost surely. Suppose $\{\mathcal{F}_m\}$ is a non-decreasing sequence of σ -algebras. Then, $E[X_m | \mathcal{F}_m] \rightarrow 0$ almost surely.*

²⁰This follows from the observation 2 must choose the announcement $p \in [0, 1]$ to minimize $(1 - p_i)p^2 + p_i(1 - p)^2$, which is accomplished by choosing $p = p_i$.

²¹The existence of such an equilibrium is established by Theorem 6.1 of Fudenberg and Levine (1983). In a finite horizon, existence would be ensured by Glicksberg's theorem (that any game with continuous payoff functions on compact subsets of Euclidean spaces has a (possibly mixed) equilibrium. Fudenberg and Levine use a limiting argument, exploiting discounting to achieve the required continuity, to extend this result to infinite-horizon games.

Proof. Define $Y_m = \sup_{m' \geq m} |X_{m'}|$. The sequence $\{Y_m\}$ is a non-increasing sequence of random variables converging to 0 almost surely. Since $E[Y_{m+1}|\mathcal{F}_m] \leq E[Y_m|\mathcal{F}_m]$ almost surely, $\{E[Y_m|\mathcal{F}_m]\}$ is a bounded supermartingale with respect to $\{\mathcal{F}_m\}$, and so there exists Y_∞ such that $E[Y_m|\mathcal{F}_m] \rightarrow Y_\infty$ almost surely. But since $E[E[Y_m|\mathcal{F}_m]] = E[Y_m] \rightarrow 0$, $EY_\infty = 0$. Since $Y_\infty \geq 0$ almost surely, we have $Y_\infty = 0$ almost surely.²²

Finally, $-E[Y_m|\mathcal{F}_m] \leq E[X_m|\mathcal{F}_m] \leq E[Y_m|\mathcal{F}_m]$ implies $E[X_m|\mathcal{F}_m] \rightarrow 0$ almost surely. ■

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., AND J. P. THOMAS (1995): “Reputation and Commitment in Two-Person Repeated Games Without Discounting,” *Econometrica*, 63(6), 1401–1419.
- FUDENBERG, D., D. KREPS, AND E. MASKIN (1990): “Repeated Games with Long-Run and Short-Run Players,” *Review of Economic Studies*, 57, 555–574.

²²This proof is taken from Hart (1985, Lemma 4.24), who credits J.-F. Mertens.

- FUDENBERG, D., AND D. K. LEVINE (1983): “Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games,” *Journal of Economic Theory*, 31, 227–250.
- (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.
- HART, S. (1985): “Non-Zero-Sum Two-Person Repeated Games with Incomplete Information,” *Mathematics of Operations Research*, 10(1), 117–153.
- HART, S., AND D. SCHMEIDLER (1989): “Existence of Correlated Equilibria,” *Mathematics of Operations Research*, 14(1), 18–25.
- 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.