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

PIER Working Paper 04-031

"Disappearing Private Reputations in Long-Run Relationships"

by

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

http://ssrn.com/abstract=571741

Disappearing Private Reputations in Long-Run Relationships*

Martin W. Cripps
John M. 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 28, 2004
(first version March 1, 2004)

Abstract

For games of *public* reputation with uncertainty over types and imperfect public monitoring, Cripps, Mailath, and Samuelson (2004) showed that an informed player facing short-lived uninformed opponents cannot maintain a permanent reputation for playing a strategy that is *not* part of an equilibrium of the game without uncertainty over types. This paper extends that result to games in which the uninformed player is long-lived and has private beliefs, so that the informed player's reputation is *private*. We also show that the rate at which reputations disappear is *uniform* across equilibria and that reputations disappear in sufficiently long discounted finitely-repeated games. *Journal of Economic Literature* Classification Numbers C70, C78. *Keywords*: Reputation, Imperfect Monitoring, Repeated Games, Commitment, Private Beliefs.

*We thank the audiences at many seminars for their helpful comments, and the National Science Foundation (grants SES-0095768, SES-9911219, SES-0241506, and SES-0350969) and the Russell Sage Foundation (82-02-04) for financial support. We also thank the Studienzentrum Gerzensee and ESSET 2004, whose hospitality facilitated Sections 3.3 and 3.4.

1 Introduction

Reputation games capture settings in which a long-lived player benefits from the perception that her characteristics may be different than they actually are. Reputation effects arise most cleanly when a long-lived player faces a sequence of short-lived players who believe the long-lived player might be committed to the stage-game "Stackelberg" action. In such a setting, the Stackelberg payoff provides a lower bound on the long-lived player's average payoff, provided she is sufficiently patient (Fudenberg and Levine (1989), Fudenberg and Levine (1992)).

In an earlier paper (Cripps, Mailath, and Samuelson (2004)), we showed that if monitoring is imperfect and the reputation of the long-lived player is *public*, meaning that the public signals allow the long-lived player to infer the short-lived players' beliefs about the long-lived player's type, then reputation effects eventually disappear. Almost surely, the short-lived player eventually learns the type of the long-lived player.

Many long-run relationships involve two (or more) long-lived players. Reputation effects arise in this setting as well, and can be more powerful than when the uninformed player is short-lived. Intertemporal incentives can induce the uninformed agent to choose actions even more advantageous to the informed long-lived player than the myopic best reply to the Stackelberg action (Celentani, Fudenberg, Levine, and Pesendorfer (1996)). In addition, it is natural for an analysis of long-lived uninformed players to encompass *private* reputations: the actions of both players are not only imperfectly monitored, but the monitoring need not have the special structure required for the informed player to infer the uninformed player's beliefs. Instead, the uninformed player's beliefs can depend critically on his own past actions, which the informed player cannot observe.[1]

In this paper, we show that reputations eventually disappear when the uninformed player is long-lived and beliefs are private.[2] We also improve on our

[1] For example, the inferences a firm draws from market prices may depend upon the firm's output choices, which others do not observe. Because private reputations arise when the uninformed player privately observes his *own* past actions, they occur most naturally with a single, long-lived uninformed player rather than a sequence of short-lived players. In Cripps, Mailath, and Samuelson (2004), we assumed that the short-run player's actions are public, allowing a natural interpretation of the assumption that short-run players' observed their predecessors' actions, but also ensuring that player 1's reputation (player 2's belief) is public.

[2] Cripps, Mailath, and Samuelson (2004, Theorem 6) is a *partial* result for the case of a long-lived uninformed player whose beliefs are *public*. That result is unsatisfactory, even for the public-reputation case, in that it imposes a condition on the behavior of the long-lived uninformed player *in equilibrium*. See footnote 5 for more details.

earlier paper by showing that the rate at which reputations disappear is *uniform* across equilibria (Theorem 3), and that reputations disappear in sufficiently long discounted finitely-repeated games (Theorem 4).

In our analysis, the long-lived informed player (player 1) may be a commitment type that plays an exogenously specified strategy or a normal type that maximizes expected payoffs. We show 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 repeated game, almost surely the uninformed player (player 2) will learn that a normal long-lived player is indeed normal. Thus, a long-lived player cannot indefinitely maintain a reputation for behavior that is not credible given her type.

Establishing such a result for the case of public reputations and short-lived uninformed players is relatively straightforward (Cripps, Mailath, and Samuelson (2004)). Since monitoring is imperfect, deviations from equilibrium play by player 1 cannot be unambiguously detected by player 2, precluding the trigger-strategy equilibria that support permanent reputations in perfect-monitoring games. Instead, the long-run convergence of beliefs ensures that eventually *any* current signal of play has an arbitrarily small effect on player 2's beliefs. Thus, when reputations are public, player 1 eventually *knows* that player 2's beliefs have nearly converged and hence that playing differently from the commitment strategy will incur virtually no cost in terms of altered beliefs. Coupled with discounting, this ensures that deviations from the commitment strategy have virtually no effect on the payoffs from continuation play. But the long-run effect of many such deviations from the commitment strategy would be to drive the equilibrium to full revelation. Public reputations can thus be maintained only if the gains from deviating from the commitment strategy are arbitrarily small, 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.[3]

The situation is more complicated in the private-reputation case, where player 2's beliefs are *not* known by player 1. Now, player 1 may not know when deviations from the commitment strategy have relatively little effect on beliefs and hence are relatively costless. Making the leap from the preceding intuition to our main result thus requires showing that there is a set of histories under which player 2's beliefs have nearly converged, *and* under which player 1 is eventually

[3]This argument does not carry over to repeated games without discounting, where small changes in beliefs, with implications only for distant behavior, can still have large payoff implications.

relatively certain player 2 has such beliefs.

In general, one cannot expect player 1's beliefs about player 2's beliefs to be very accurate when the latter depend on private histories. A key step in our proof is to show that whenever player 2's private history induces him to act as if he is convinced of some important characteristic of player 1, eventually player 1 must become convinced that such a private history did indeed occur (Lemma 3). In particular, if this private history ensured that player 2 is almost convinced that he faces a commitment type, and acts on this belief, then this eventually becomes known to player 1.

As in the case where player 1's reputation is public, the impermanence of reputation also arises at the behavioral level. Asymptotically, continuation play in every Nash equilibrium is a correlated equilibrium of the complete-information game (Theorem 5). While the set of Nash equilibrium payoffs in the game with complete information is potentially very large when player 2 is sufficiently patient (suggesting that limiting behavior to that set imposes few restrictions), we emphasize that our analysis holds for *all* degrees of patience of the players. When player 2 is impatient, as in the extreme case of short-run player 2s, reputations can ensure payoffs for player 1 that cannot be obtained under complete information. Our result (that limiting behavior must be consistent with complete information) shows that this effect is transitory.

More importantly, reputation arguments are also of interest for their ability to restrict, rather than expand, the set of equilibrium outcomes. For example, reputation arguments are important in perfect-monitoring games with patient players, precisely because they impose tight bounds on (rather than expanding) the set of equilibrium payoffs. Our results caution that one cannot assume that such selection effects are long-lasting.

For expositional clarity, this paper considers a long-lived informed player who can be one of two possible types—a commitment and a normal type—facing a single long-lived uninformed player, in a game of imperfect public monitoring. The argument of Cripps, Mailath, and Samuelson (2004, Section 6.1) can be used to extend our results to many possible commitment types. The final section of this paper explains how our results can be extended to the case of private monitoring (where reputations are necessarily private).

Our analysis subsumes a private-reputation model with a sequence of short-lived uninformed players. In several places, the arguments for the latter case are simpler and considerably more revealing, primarily because we can then restrict attention to simpler commitment strategies. Accordingly, where appropriate, we give the simpler argument for short-lived uninformed players as well as the more involved argument for the long-lived uninformed player.

2 The Complete-Information Game

We begin with an infinitely repeated game with imperfect public monitoring. The stage game is a two-player simultaneous-move finite game of public monitoring. Player 1 chooses an action $i \in \{1, 2, ..., I\} \equiv I$ and player 2 chooses an action $j \in \{1, 2, ..., J\} \equiv J$. The public signal, y, is drawn from the 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 $\pi_2(i, j) = \sum_y f_2(j, y)\, \rho_{ij}^y$.

We assume the public signals have full support (Assumption 1), so every signal y is possible after any action profile. We also assume that with sufficiently many observations, either player can correctly identify, from the frequencies of the signals, any fixed stage-game action of their opponent (Assumptions 2 and 3).

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

Assumption 2 (IDENTIFICATION OF 1) *For all* $j \in J$*, the* I *columns in the matrix* $(\rho_{ij}^y)_{y \in Y, i \in I}$ *are linearly independent.*

Assumption 3 (IDENTIFICATION OF 2) *For all* $i \in I$*, the* J *columns in the matrix* $(\rho_{ij}^y)_{y \in Y, j \in J}$ *are linearly independent.*

The stage game is infinitely repeated. Player 1 ("she") is a long-lived player with discount factor $\delta_1 < 1$. Player 2 ("he") is either short-lived, in which case a new player 2 appears in each period, or is also long-lived, in which case player 2's discount factor δ_2 may differ from δ_1. Each player observes the realizations of the public signal and his or her own past actions. (If player 2 is short-lived, he observes the actions chosen by the previous player 2's). Player 1 in period t thus has a *private history*, consisting of the public signals and her own past actions, denoted by $h_{1t} \equiv ((i_0, y_0), (i_1, y_1), \ldots, (i_{t-1}, y_{t-1})) \in H_{1t} \equiv (I \times Y)^t$. Similarly, a *private history* for player 2 is denoted $h_{2t} \equiv ((j_0, y_0), (j_1, y_1), \ldots, (j_{t-1}, y_{t-1})) \in H_{2t} \equiv (J \times Y)^t$. The *public history* observed by both players is the sequence $(y_0, y_1, \ldots, y_{t-1}) \in Y^t$. The filtration on $(I \times J \times Y)^\infty$ induced by the private histories of player $\ell = 1, 2$ is denoted $\{\mathcal{H}_{\ell t}\}_{t=0}^\infty$, while the filtration induced by the public histories $(y_0, y_1, ..., y_{t-1})$ is denoted $\{\mathcal{H}_t\}_{t=0}^\infty$.

In Cripps, Mailath, and Samuelson (2004), we assumed that the public signal included player 2's action. This ensures that player 1 knows everything player 2 does, including player 2's beliefs. Here, only player 2 observes his action, breaking the link between 2's beliefs and 1's beliefs about those beliefs.

The long-lived players' payoffs in the infinite horizon game are the averaged discounted sum of stage-game payoffs, $(1-\delta_\ell)\sum_{\tau=0}^{\infty}\delta_\ell^\tau \pi_\ell(i_\tau, j_\tau)$ for $\ell = 1, 2$. The random variable $\pi_{\ell t}$ denotes average discounted payoffs in period t,

$$\pi_{\ell t} \equiv (1-\delta_\ell)\sum_{\tau=t}^{\infty}\delta_\ell^{\tau-t}\pi_\ell(i_\tau, j_\tau). \tag{1}$$

If player 2 is short-lived, the period-t player 2 has payoffs $\pi_2(i_t, j_t)$.

A behavior strategy for player 1 (respectively, 2) is a map, $\sigma_1 : \cup_{t=0}^{\infty} H_{1t} \to \Delta^I$ (resp., $\sigma_2 : \cup_{t=0}^{\infty} H_{2t} \to \Delta^J$), from all private histories to the set of distributions over current actions. For $\ell = 1, 2$, σ_ℓ defines a sequence of functions $\{\sigma_{\ell t}\}_{t=0}^{\infty}$ with $\sigma_{1t} : H_{1t} \to \Delta^I$ and $\sigma_{2t} : H_{2t} \to \Delta^J$. Each function $\sigma_{\ell t}$ denotes the t^{th} period behavior strategy of σ_ℓ. The strategy profile $\sigma = (\sigma_1, \sigma_2)$ induces a probability distribution P^σ over $(I \times J \times Y)^\infty$. Let $E^\sigma[\,\cdot \mid \mathcal{H}_{\ell t}]$ denote player ℓ's expectations with respect to this distribution conditional on $\mathcal{H}_{\ell t}$.

A Nash equilibrium for the case of two long-lived players requires player ℓ's strategy to maximize the expected value of $\pi_{\ell 0}$, the discounted value of payoffs in period zero:

Definition 1 *A* Nash equilibrium of the complete-information game with a long-lived player 2 *is a strategy profile* $\sigma = (\sigma_1, \sigma_2)$ *such that* $E^\sigma[\pi_{10}] \geq E^{(\sigma_1', \sigma_2)}[\pi_{10}]$ *for all* σ_1' *and* $E^\sigma[\pi_{20}] \geq E^{(\sigma_1, \sigma_2')}[\pi_{20}]$ *for all* σ_2'.

This requires that under the equilibrium profile, player ℓ's strategy maximizes continuation expected utility after any positive-probability history. For example, for player 1, $E^\sigma[\pi_{1t}|\mathcal{H}_{1t}] \geq E^{(\sigma_1', \sigma_2)}[\pi_{1t}|\mathcal{H}_{1t}]$ P^σ-almost surely for all σ_1' and all t. The assumption of full-support monitoring ensures that all histories of public signals occur with positive probability, and hence must be followed by optimal behavior in any Nash equilibrium (with long-lived or short-lived player 2's, and complete or incomplete information). Consequently, any Nash equilibrium outcome is also the outcome of a perfect Bayesian equilibrium.

For future reference, when player 2 is long-lived,

$$BR^L(\sigma_1) \equiv \{\sigma_2 : E^\sigma[\pi_{20}] \geq E^{(\sigma_1, \sigma_2')}[\pi_{20}]\ \forall \sigma_2'\}$$

is the set of player 2's best replies to σ_1 in the game with complete information.

When player 2 is short-lived, in equilibrium, player 2 plays a best response after every equilibrium history. Player 2's strategy σ_2 is then a best response to σ_1 if, for all t,

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

Denote the set of such best responses by $BR^S(\sigma_1)$. The definition of a Nash equilibrium for this case is:

Definition 2 *A* Nash equilibrium of the complete-information game with a short-lived player 2 *is a strategy profile* $\sigma = (\sigma_1, \sigma_2)$ *such that* $E^{\sigma}[\pi_{10}] \geq E^{(\sigma_1', \sigma_2)}[\pi_{10}]$ *for all* σ_1' *and* $\sigma_2 \in BR^S(\sigma_1)$.

3 The Incomplete-Information Game: Disappearing Reputations

We now perturb the complete-information game by introducing incomplete information about the type of player 1. At time $t = -1$, Nature selects a type of player 1. With probability $1 - p_0 > 0$, she is the "normal" type, denoted by n and with the preferences described above, who plays a repeated game strategy $\tilde{\sigma}_1$. With probability $p_0 > 0$, she is a "commitment" type, denoted by c, who plays the repeated game strategy $\hat{\sigma}_1$.

A state of the world in the incomplete information game, ω, is a type for player 1 and a sequence of actions and signals. The set of states is $\Omega \equiv \{n, c\} \times (I \times J \times Y)^{\infty}$. The prior p_0, the commitment strategy $\hat{\sigma}_1$, and the strategy profile $\tilde{\sigma} = (\tilde{\sigma}_1, \sigma_2)$, jointly induce a probability measure P on Ω, which describes how an uninformed player expects play to evolve. The strategy profile $\hat{\sigma} = (\hat{\sigma}_1, \sigma_2)$ (respectively, $\tilde{\sigma} = (\tilde{\sigma}_1, \sigma_2)$) determines a probability measure $\hat{P}$ (resp., $\tilde{P}$) on Ω, which describes how play evolves when player 1 is the commitment (resp., normal) type. Since $\tilde{P}$ and $\hat{P}$ are absolutely continuous with respect to P, any statement that holds P-almost surely, also holds $\tilde{P}$- and $\hat{P}$-almost surely. We use $E^{(\tilde{\sigma}_1, \hat{\sigma}_1, \sigma_2)}[\,\cdot\,]$ to denote expectations taken with respect to the measure P. This will usually be abbreviated to $E[\,\cdot\,]$ except where it is important to emphasize the dependence on the strategies. Also, where appropriate, we use $\tilde{E}[\,\cdot\,]$ and $\hat{E}[\,\cdot\,]$ to denote the expectations taken with respect to $\tilde{P}$ and $\hat{P}$ (instead of $E^{(\tilde{\sigma}_1, \sigma_2)}[\,\cdot\,]$ and $E^{(\hat{\sigma}_1, \sigma_2)}[\,\cdot\,]$). The filtrations $\{\mathcal{H}_{\ell t}\}_{t=0}^{\infty}$ and $\{\mathcal{H}_t\}_{t=0}^{\infty}$ will be viewed as filtrations on Ω in the obvious way.

The normal type of player 1 has the same objective function as in the complete-information game. Player 2, on the other hand, uses the information he has acquired from his time t private history to update his beliefs about player 1's type and actions, and then maximizes expected payoffs. Player 2's posterior belief in period t that player 1 is the commitment type is the $\mathcal{H}_{2t}$-measurable random variable $P(c|\mathcal{H}_{2t}) \equiv p_t : \Omega \to [0,1]$. By Assumption 1, Bayes' rule determines this posterior after all histories. 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. It therefore converges P-almost surely (and hence $\tilde{P}$- and $\hat{P}$-almost surely) to a random variable p_∞ defined on Ω. Furthermore, at any equilibrium the posterior p_t is a $\hat{P}$-submartingale and a $\tilde{P}$-supermartingale with respect to the filtration $\{\mathcal{H}_{2t}\}_t$.

3.1 Uninformed Player is Short-Lived

When player 2 is short-lived, and we are interested in the lower bounds on player 1's ex ante payoffs that arise from the existence of "Stackelberg" commitment types (as in Fudenberg and Levine (1992)), it suffices to consider commitment types who follow "simple" strategies. Consequently, when player 2 is short-lived, we assume $\hat{\sigma}_1$ specifies the same (possibly mixed) action $\varsigma_1 \in \Delta^I$ in each period independent of history (cf. Definition 4 below).

If ς_1 is part of a stage-game equilibrium, reputations need not disappear—we need only consider an equilibrium in which the normal and commitment type both play ς_1, and player 2 plays his part of the corresponding equilibrium. We are interested in commitment types who play a strategy that is *not* part of a stage-game equilibrium:[4]

Assumption 4 (NON-CREDIBLE COMMITMENT) *Player 2 has a unique best reply to ς_1 (denoted ς_2) and $\varsigma \equiv (\varsigma_1, \varsigma_2)$ is not a stage-game Nash equilibrium.*

Since ς_2 is the unique best response to ς_1, ς_2 is pure and $BR^S(\hat{\sigma}_1)$ is the singleton $\{\hat{\sigma}_2\}$, where $\hat{\sigma}_2$ is the strategy of playing ς_2 in every period. Assumption 4 implies that $(\hat{\sigma}_1, \hat{\sigma}_2)$ is not a Nash equilibrium of the complete-information infinite horizon game.

[4]If player 2 has multiple best responses, it is possible to construct equilibria of the complete information game in which player 1 always plays ς_1 in each period, irrespective of history, even if ς_1 is not part of a stage-game equilibrium (for an example, see Cripps, Mailath, and Samuelson (2004, Section 2)).

Definition 3 *A* Nash equilibrium of the incomplete-information game with short-lived uninformed players *is a strategy profile* $(\tilde{\sigma}_1, \sigma_2)$ *such that for all* σ_1', $j \in J$ *and* $t = 0, 1, ...,$

$$\begin{aligned} \tilde{E}\left[\pi_{10}\right] &\geq E^{(\sigma_1', \sigma_2)}\left[\pi_{10}\right], \textit{ and} \\ E[\, \pi_2(i_t, j_t) \mid \mathcal{H}_{2t}] &\geq E[\, \pi_2(i_t, j) \mid \mathcal{H}_{2t}], \qquad P\text{-a.s.} \end{aligned}$$

Our main result, for short-lived uninformed players, is that reputations for non-equilibrium behavior are temporary:

Theorem 1 *Suppose the monitoring distribution* ρ *satisfies Assumptions 1, 2, and 3 and the commitment action* ς_1 *satisfies Assumption 4. In any Nash equilibrium of the incomplete-information game with short-lived uninformed players,* $p_t \to 0$ $\tilde{P}$*-almost surely.*

3.2 Uninformed Player is Long-Lived

When player 2 is long-lived, non-simple Stackelberg types may give rise to higher lower bounds on player 1's payoff than do simple types. We accordingly introduce a richer set of possible commitment types, allowing arbitrary public strategies.

Definition 4 *(1) A behavior strategy* σ_ℓ, $\ell = 1, 2$, *is* public *if it is measurable with respect to the filtration induced by the public signals,* $\{\mathcal{H}_t\}_t$.
(2) A behavior strategy σ_ℓ, $\ell = 1, 2$, *is* simple *if it is a constant function.*

A public strategy induces a mixture over actions in each period that only depends on public histories. Any pure strategy is realization equivalent to a public strategy. Simple strategies, which we associated with the commitment type in Section 3.1, play the same mixture over stage-game actions in each period, and hence are trivially public.

Allowing the commitment type to play any public strategy necessitates imposing the noncredibility requirement directly on the infinitely repeated game of complete information. Mimicking Assumption 4, we require that (*i*) player 2's best response $\hat{\sigma}_2$ be unique on the equilibrium path and (*ii*) there exists a finite time T^o such that, for every $t > T^o$, a normal player 1 would almost surely want to deviate from $\hat{\sigma}_1$, given player 2's best response. That is, there is a period-t continuation strategy for player 1 that strictly increases her utility. A strategy $\hat{\sigma}_1$ satisfying these criteria at least eventually loses its credibility, and hence is said to have "no long-run credibility."

Definition 5 *The strategy* $\hat{\sigma}_1$ *has* no long-run credibility *if there exists* T^o *and* $\varepsilon^o > 0$ *such that, for every* $t \geq T^o$,

(1) $\hat{\sigma}_2 \in BR^L(\hat{\sigma}_1)$ *implies that with* $P^{(\hat{\sigma}_1,\hat{\sigma}_2)}$*-probability one,* $\hat{\sigma}_{2t}$ *is pure and*

$$E^{\hat{\sigma}}[\,\pi_{2t} \mid \mathcal{H}_{2t}\,] > E^{(\hat{\sigma}_1,\sigma_2')}[\,\pi_{2t} \mid \mathcal{H}_{2t}\,] + \varepsilon^o,$$

for all σ_2' *attaching probability zero to the action played by* $\hat{\sigma}_{2t}(h_{2t})$ *after* $P^{(\hat{\sigma}_1,\hat{\sigma}_2)}$*-almost all* $h_{2t} \in H_{2t}$*, and*

(2) *there exists* $\tilde{\sigma}_1$ *such that, for* $\hat{\sigma}_2 \in BR^L(\hat{\sigma}_1)$, $P^{(\hat{\sigma}_1,\hat{\sigma}_2)}$*-almost surely,*

$$E^{(\tilde{\sigma}_1,\hat{\sigma}_2)}[\,\pi_{1t} \mid \mathcal{H}_{1t}\,] > E^{\hat{\sigma}}[\,\pi_{1t} \mid \mathcal{H}_{1t}\,] + \varepsilon^o.$$

This definition captures the two main features of Assumption 4, a unique best response and absence of equilibrium, in a dynamic setting. In particular, the stage-game action of any simple strategy satisfying Definition 5 satisfies Assumption 4. In assuming the best response is unique, we need to avoid the possibility that there are multiple best responses to the commitment action "in the limit" (as t gets large). We do so by imposing a uniformity condition in Definition 5.1, that inferior responses reduce payoffs by at least ε^o. The condition on the absence of equilibrium in Definition 5.2 similarly ensures that for all large t, player 1 can strictly improve on the commitment action. Again it is necessary to impose uniformity to avoid the possibility of an equilibrium in the limit.[5]

Any $\hat{\sigma}_1$ that does *not* satisfy Definition 5 must have (at least in the limit) periods and histories where, given player 2 is best responding to $\hat{\sigma}_1$, player 1 prefers to stick to her commitment. In other words, $\hat{\sigma}_1$ is a credible commitment, in the limit, at least some of the time.

Equilibrium when the uninformed player is long-lived is:

Definition 6 *A* Nash equilibrium of the incomplete-information game with a long-lived uninformed player *is a strategy profile* $(\tilde{\sigma}_1, \sigma_2)$ *such that,*

$$\begin{aligned}\tilde{E}\,[\pi_{10}] &\geq E^{(\sigma_1',\sigma_2)}\,[\pi_{10}]\;,\;\; \forall \sigma_1', \text{ and}\\ E[\pi_{20}] &\geq E^{(\tilde{\sigma}_1,\hat{\sigma}_1,\sigma_2')}[\pi_{20}],\;\; \forall \sigma_2'.\end{aligned}$$

[5] Cripps, Mailath, and Samuelson (2004) show that reputations disappear when the commitment strategy satisfies the second, but not necessarily the first, condition (such a strategy was said to be *never an equilibrium strategy in the long run*). However, that result also requires the commitment strategy to be implementable by a finite automaton, and more problematically, the result itself imposed a condition on the behavior of player 2 in the equilibrium of the game with incomplete information. We do neither here. Consequently, unlike our earlier paper, the long-lived player result implies the result for short-lived players.

Our result for games where player 2 is long-lived, which implies Theorem 1, is:

Theorem 2 *Suppose ρ satisfies Assumptions 1, 2, and 3, and that the commitment type's strategy $\hat{\sigma}_1$ is public and has no long run credibility. Then in any Nash equilibrium of the game with incomplete information, $p_t \to 0$ $\tilde{P}$-almost surely.*

We have followed the standard practice of working with commitment types whose behavior is fixed. If we instead modeled commitment types as strategic agents whose payoffs differed from those of normal types, we would obtain the following: Under Assumptions 1–3, in any Nash equilibrium in which the "commitment-payoff" type plays a public strategy with no long run credibility for the "normal-payoff" type, $p_t \to 0$ $\tilde{P}$-almost surely.

3.3 Uniform Disappearance of Reputations

Theorem 2 leaves open the possibility that while reputations do asymptotically disappear in every equilibrium, for any period T, there may be equilibria in which reputations survive beyond T. We show here that that possibility cannot arise: there is some T after which reputations have disappeared in *all* Nash equilibria. Intuitively, a sequence of Nash equilibria with reputations persisting beyond period $T \to \infty$ implies the (contradictory) existence of a limiting Nash equilibrium with a permanent reputation.

Theorem 3 *Suppose ρ satisfies Assumptions 1, 2, and 3, and that the commitment type's strategy $\hat{\sigma}_1$ is public and has no long run credibility. For all $\varepsilon > 0$, there exists T, such that for all Nash equilibria, σ, of the game with incomplete information,*

$$\tilde{P}^{\sigma}(p_t^{\sigma} < \varepsilon,\ \forall t > T) > 1 - \varepsilon,$$

where $\tilde{P}^{\sigma}$ is the probability measure induced on Ω by σ and the normal type, and p_t^{σ} is the associated posterior of player 2 on the commitment type.

Proof. Suppose not. Then there exists $\varepsilon > 0$ such that for all T, there is a Nash equilibrium σ^T such that

$$\tilde{P}^{T}(p_t^{T} < \varepsilon,\ \forall t > T) \leq 1 - \varepsilon,$$

where $\tilde{P}^T$ is the measure induced by the normal type under σ^T and p_t^T is the posterior in period t under σ^T.

Since the space of strategy profiles is sequentially compact in the product topology, there is a convergent subsequence $\{\sigma^{T_k}\}$, with limit σ^*. We can relabel this sequence so that $\sigma^k \to \sigma^*$ and

$$\tilde{P}^k(p_t^k < \varepsilon,\ \forall t > k) \leq 1 - \varepsilon,$$

i.e.,

$$\tilde{P}^k(p_t^k \geq \varepsilon \text{ for some } t > k) \geq \varepsilon.$$

Since each σ^k is a Nash equilibrium, $p_t^k \to 0$ $\tilde{P}^k$-a.s. (Theorem 2), and so there exists K_k such that

$$\tilde{P}^k(p_t^k < \varepsilon,\ \forall t \geq K_k) \leq 1 - \varepsilon/2.$$

Consequently, for all k,

$$\tilde{P}^k(p_t^k \geq \varepsilon,\ \text{for some } t,\ k < t < K_k) \geq \varepsilon/2.$$

Let τ_k denote the stopping time

$$\tau_k = \min\{t > k : p_t^k \geq \varepsilon\},$$

and q_t^k the associated stopped process,

$$q_t^k = \begin{cases} p_t^k, & \text{if } t < \tau_k, \\ \varepsilon, & \text{if } t \geq \tau_k. \end{cases}$$

Note that q_t^k is a supermartingale under $\tilde{P}^k$ and that for $t < k$, $q_t^k = p_t^k$.

Observe that for all k and $t \geq K_k$,

$$\tilde{E}q_t^k \geq \varepsilon \tilde{P}^k(\tau_k \leq t) \geq \varepsilon^2/2.$$

Since σ^* is a Nash equilibrium, $p_t^* \to 0$ $\tilde{P}^*$-a.s. (appealing to Theorem 2 again), and so there exists a date s such that

$$\tilde{P}^*(p_s^* < \varepsilon^2/12) > 1 - \varepsilon^2/12.$$

Then,

$$\tilde{E}^* p_s^* \leq \frac{\varepsilon^2}{12}(1 - \frac{\varepsilon^2}{12}) + \frac{\varepsilon^2}{12} < \frac{\varepsilon^2}{6}.$$

Since $\sigma^k \to \sigma^*$ in the product topology, there is a $k' > s$ such that for all $k \geq k'$,

$$\tilde{E}^k p_s^k < \frac{\varepsilon^2}{3}.$$

But since $k' > s$, $q_s^k = p_s^k$ for $k \geq k'$ and so for any $t \geq K_k$,

$$\begin{aligned} \frac{\varepsilon^2}{3} &> \tilde{E}^k p_s^k = \tilde{E}^k q_s^k \\ &\geq \tilde{E}^k q_t^k \geq \frac{\varepsilon^2}{2}, \end{aligned} \tag{2}$$

which is a contradiction. ■

3.4 Disappearing reputations in discounted finitely-repeated games

In this section we show that reputations also disappear in sufficiently long discounted finitely-repeated games of incomplete information. We first describe the finitely repeated game with incomplete information. If the commitment type plays a simple strategy of playing ς_1 in every period, with ς_1 satisfying Assumption 4, then the description of the finitely repeated game for differing repetitions is straightforward: The commitment type plays ς_1 in every period. More generally, if $\hat{\sigma}_1^T$ is the commitment type's strategy in the T-period game, we require that the sequence $\{\hat{\sigma}_1^T\}$ converge to a strategy $\hat{\sigma}_1$ of the infinitely repeated game that has no long-run credibility.

Theorem 4 *Suppose ρ satisfies Assumptions 1, 2, and 3, and $\hat{\sigma}_1$ is a public strategy of the infinitely repeated game with no long run credibility. Let G^T denote the T-period repeated game of incomplete information in which the commitment type plays according to $\hat{\sigma}_1^T$. Suppose for all t, $\hat{\sigma}_{1t}^T \to \hat{\sigma}_{1t}$ as $T \to \infty$. For all $\varepsilon > 0$, there exists T such that for all $T' \geq T$ and for all Nash equilibria σ of $G^{T'}$,*

$$\tilde{P}^\sigma(p_t^\sigma < \varepsilon, \, \forall t \geq T) > 1 - \varepsilon,$$

where $\tilde{P}^\sigma$ is the probability measure induced on $(I \times J \times Y)^{T'}$ by σ and the normal type, and p_t^σ is the associated posterior of player 2 on the commitment type.

Proof. Suppose not. Then there exists $\varepsilon > 0$, such that for all T, there exists $T' \geq T$ and a Nash equilibrium σ^T of the T'-period finitely repeated game with

$$\tilde{P}^T(p_t^T < \varepsilon, \, \forall t \geq T) \leq 1 - \varepsilon,$$

where $\tilde{P}^T$ is the probability measure induced in the T'-period repeated game by σ^T and the normal type, and p_t^T is the associated posterior.

A standard diagonalization argument yields a subsequence $\{\sigma^{T_k}\}$ and a strategy profile in the infinitely repeated game, σ^*, with the property that for all t, $\sigma_{\ell t}^{T_k} \to \sigma_{\ell t}^*$ for $\ell = 1, 2$.[6] Moreover, since each σ^{T_k} is a Nash equilibrium of increasingly long finitely repeated games and $\hat{\sigma}_{1t}^{T_k} \to \hat{\sigma}_{1t}$, σ^* is a Nash equilibrium of the infinitely repeated game with incomplete information in which the commitment type plays $\hat{\sigma}_1$. We can relabel this sequence so that $\sigma_t^k \to \sigma_t^*$ for each t and

$$\tilde{P}^k(p_t^k < \varepsilon,\ \forall t > k) \leq 1 - \varepsilon.$$

Letting T_k be the length of the finitely repeated game corresponding to σ^k, we have (recall that the initial period is period 0)

$$\tilde{P}^k(p_t^k \geq \varepsilon,\ \text{for some } t,\ k < t < T_k) \geq \varepsilon.$$

The proof now proceeds as that of Theorem 3, with (2) evaluated at $t = T_k - 1$.

■

3.5 Asymptotic Equilibrium Play

The impermanence of reputations has implications for behavior as well as beliefs. 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' beliefs about how play will continue vanish in the limit. This is stronger than the convergence to *subjective* equilibria obtained by Kalai and Lehrer (1995, Corollary 4.4.1),[7] though with stronger assumptions.

We present the result for the case of a long-run player 2, since only straightforward modifications are required (imposing the appropriate optimality conditions period-by-period) to address short-run player 2's. To begin, we describe some notation for the correlated equilibrium of the repeated game with imperfect monitoring. We use the term *period-t continuation game* for the game with initial period in period t.[8] We use the notation $t' = 0, 1, 2, ...$ for a period of play in a

---

[6]For each t, $\sigma_t^{T_k}$ and σ_t^* are elements of the same finite dimensional Euclidean space.

[7]In a subjective correlated equilibrium, the measure in (3) can differ from the measure in (4).

[8]Since 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.

continuation game (which may be the original game) and t for the time elapsed prior to the start of the period-t continuation game. A pure strategy for player 1, s_1, is a sequence of maps $s_{1t'} : H_{1t'} \to I$ for $t' = 0, 1, \ldots$.[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

$$\begin{aligned} u_1(s_1, s_2) &\equiv E^{(s_1,s_2)}[\,\pi_{10}\,], \text{ and} \\ u_2(s_1, s_2) &\equiv E^{(s_1,s_2)}[\,\pi_{20}\,]. \end{aligned}$$

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.

We follow Hart and Schmeidler (1989) in using the ex ante definition of correlated equilibria for infinite pure-strategy sets:

Definition 7 *A* correlated equilibrium *of the complete-information game is a measure μ on $(S, \mathcal{S})$ such that for all $\mathcal{S}$-measurable functions $\zeta_1 : S_1 \to S_1$ and $\zeta_2 : S_2 \to S_2$,*

$$\int_S [u_1(s_1, s_2) - u_1(\zeta_1(s_1), s_2)]d\mu \;\geq\; 0, \textit{ and} \tag{3}$$

$$\int_S [u_2(s_1, s_2) - u_2(s_1, \zeta_2(s_2))]d\mu \;\geq\; 0. \tag{4}$$

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 $\tau \geq 0$, we have

$$\mu_n\big|_{I^{(I\times Y)^\tau}\times J^{(J\times Y)^\tau}} \to \mu\big|_{I^{(I\times Y)^\tau}\times J^{(J\times Y)^\tau}}.$$

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 payoffs are discounted, the product topology is strong enough to guarantee continuity of $u_\ell : \mathcal{M} \to \mathbb{R}$. The set of mixed strategies for player ℓ is denoted by $\mathcal{M}_\ell$.

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

[9] Recall that σ_1 denotes general behavior strategies.

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 period-t continuation game. This behavior strategy is realization equivalent to a mixed strategy $\tilde{\lambda}^{h_{1t}} \in \mathcal{M}_1$ for the period-t 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 $\zeta_1 : S_1 \to S_1$,

$$\int_S u_1(s_1, s_2) d\tilde{\mu}_t \geq \int_S u_1(\zeta_1(s_1), s_2) d\tilde{\mu}_t. \tag{5}$$

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 $\xi_2 : S_2 \times [0, 1] \to S_2$,

$$\int_{S\times[0,1]} u_2(s_1, s_2) d(p_0\hat{\rho}_t+(1-p_0)\tilde{\rho}_t) \geq \int_{S\times[0,1]} u_2(s_1, \xi_2(s_2, p_t)) d(p_0\hat{\rho}_t+(1-p_0)\tilde{\rho}_t). \tag{6}$$

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 equilibrium 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 fact that $\mathcal{M}$ is sequentially compact, and hence every subsequence of $\{\tilde{\mu}_t\}$ has a convergent subsubsequence. The proof of the following is in the Appendix:

Theorem 5 *Fix a Nash equilibrium of the incomplete-information game and suppose $p_t \to 0$ $\tilde{P}$-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 of $\{\tilde{\mu}_t\}$ 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.[10]

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 (with short-run players) 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 $\delta_1 > \bar{\delta}$, 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 5 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 $\bar{\delta}(p_0) \to 1$ as $p_0 \to 0$), we fix δ_1 and examine asymptotic play, so that eventually p_t is sufficiently small that $\delta_1 < \bar{\delta}(p_t)$.

4 Proofs of Theorems 1 and 2

The short-lived uninformed player case is a special case of the long-lived player case. However, the proof for the long-lived uninformed player is quite complicated, while the short-lived player case illustrates many of the issues in a simpler setting. In what follows, references to the incomplete information game without further qualification refer to the game with the long-lived uninformed player, and so the discussion also covers short-lived uninformed players (where $\hat{\sigma}_1(h_s) = \varsigma_1$ for all h_s). Whenever it is helpful, however, we also give informative simpler arguments for the case of short-lived uninformed players.

The basic strategy of our proof is to show that if player 2 is not eventually convinced that player 1 is normal, then he must be convinced that player 1 is playing like the commitment type (Lemma 1) and hence player 2 plays a best response to the latter. Our earlier paper proceeded by arguing that the normal type

[10] 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. As far as we are aware, 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?

then has an incentive to deviate from the commitment strategy (since the latter has no long-run credibility), which forms the basis for a contradiction (with player 2's belief that the two types of player 1 are playing identically). The difficulty in applying this argument in our current setting is that player 1 needs to know player 2's private history h_{2t} in order to predict 2's period-t beliefs and hence behavior. Unfortunately, player 1 knows only her own private history h_{1t}. Our argument thus requires showing that player 1 eventually "almost" knows the relevant features of player 2's history.

4.1 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 normal type is, in the limit, identical to his expectation of the strategy played by the commitment type, *or* player 2's posterior probability that player 1 is the commitment type converges to zero (given that player 1 is indeed normal). This is an extension of a familiar merging-style argument to the case of imperfect monitoring. If, 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.

The proof of Lemma 1 in Cripps, Mailath, and Samuelson (2004) applies to the current setting without change:

Lemma 1 *Suppose Assumptions 1 and 2 are satisfied and $\hat{\sigma}_1$ is public. In any Nash equilibrium of the game with incomplete information,*[11]

$$\lim_{t\to\infty} p_t(1-p_t)\left\|\hat{\sigma}_{1t} - \tilde{E}[\,\tilde{\sigma}_{1t} \mid \mathcal{H}_{2t}\,]\right\| = 0, \qquad P\text{-a.s.} \tag{7}$$

Condition (7) says that almost surely either player 2's best prediction of the normal type's behavior at the current stage is arbitrarily close to his best prediction of the commitment type's behavior (that is, $\|\hat{\sigma}_{1t} - \tilde{E}[\,\tilde{\sigma}_{1t} \mid \mathcal{H}_{2t}\,]\,\| \to 0$), or the type is revealed (that is, $p_t(1-p_t) \to 0$). However, $\lim p_t < 1$ $\tilde{P}$-almost surely, and hence (7) implies a simple corollary:

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

Corollary 1 *Suppose Assumptions 1 and 2 are satisfied and $\hat{\sigma}_1$ is public. In any Nash equilibrium of the game with incomplete information,*

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

4.2 Player 2's Beliefs about his Future Behavior

We now examine the consequences of the existence of a $\tilde{P}$-positive measure set of states on which reputations do not disappear, i.e., $\lim_{t\to\infty} p_t(\omega) > 0$. The normal and the commitment types eventually play the same strategy on these states (Lemma 1). Consequently, we can show that on a positive probability subset of these states, player 2 eventually attaches high probability to the event that in all future periods he will play a best response to the commitment type.

As $\hat{\sigma}_1$ is public, player 2 has a best response to $\hat{\sigma}_1$ that is also public. Moreover, this best response is unique on the equilibrium path for all $t > T^o$ (by Definition 5). We let $j^*(h_t)$ denote the action that is the pure best-response after the public history h_t, for all $t > T^o$. Note that $j^*(h_t)$ is $\mathcal{H}_t$-measurable. The event that player 2 plays a best response to the commitment strategy in all periods after $t > T^o$ is then defined as

$$G_t^o \equiv \{\, \omega : \sigma_{2s}^{j^*(h_s(\omega))}(h_{2s}(\omega)) = 1, \forall s \geq t \,\},$$

where $h_s(\omega)$ (respectively, $h_{2s}(\omega)$) is the s-period public (resp., 2's private) history of ω.

When the uninformed players are short-lived, $\hat{\sigma}_1$ is simple and player 2 has a unique best reply, $BR^S(\varsigma_1) = \{\varsigma_2\}$, so

$$G_t^o = \{\omega : \sigma_{2s}(h_{2s}(\omega)) = \varsigma_2, \forall s \geq t\}\,.$$

With this in hand we can show that if player 2 does not eventually learn that player 1 is normal, then he eventually attaches high probability to thereafter playing a best response to the commitment type:

Lemma 2 *Suppose the hypotheses of Theorem 2 hold,*[12] *and suppose there is a Nash equilibrium in which reputations do not necessarily disappear, i.e., $\tilde{P}(A) > 0$, where $A \equiv \{p_t \nrightarrow 0\}$. There exists $\eta > 0$ and $F \subset A$, with $\tilde{P}(F) > 0$, such that, for any $\xi > 0$, there exists T for which, on F,*

$$p_t > \eta, \qquad \forall t \geq T,$$

[12]This lemma does not require Assumption 3.

and

$$\tilde{P}\left(G_t^o \mid \mathcal{H}_{2t}\right) > 1 - \xi, \qquad \forall t \geq T. \tag{8}$$

Proof. Since $\tilde{P}(A) > 0$ and p_t converges almost surely, there exists $\mu > 0$ and $\eta > 0$ such that $\tilde{P}(D) > 2\mu$, where $D \equiv \{\omega : \lim_{t\to\infty} p_t(\omega) > 2\eta\}$. The random variables $\|\hat{\sigma}_{1s} - \tilde{E}[\tilde{\sigma}_{1t}|\mathcal{H}_{2t}]\|$ tend $\tilde{P}$-almost surely to zero on D (by Corollary 1). Consequently, the random variables $Z_t \equiv \sup_{s\geq t} \|\hat{\sigma}_{1s} - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_{2s}]\|$ also converge $\tilde{P}$-almost surely to zero on D. Thus, from Hart (1985, Lemma 4.24), $\tilde{E}[\mathbf{1}_D Z_t \mid \mathcal{H}_{2t}]$ converge almost surely to zero, where $\mathbf{1}_D$ is the indicator for the event D. Define $A_t \equiv \{\omega : \tilde{E}[\mathbf{1}_D \mid \mathcal{H}_{2t}](\omega) > \frac{1}{2}\}$. The $\mathcal{H}_{2t}$-measurable event A_t approximates D (because player 2 knows his own beliefs, the random variables $d_t \equiv |\mathbf{1}_D - \mathbf{1}_{A_t}|$ converge $\tilde{P}$-almost surely to zero). Hence

$$\begin{aligned} \mathbf{1}_D \tilde{E}[Z_t \mid \mathcal{H}_{2t}] &\leq \mathbf{1}_{A_t}\tilde{E}[Z_t \mid \mathcal{H}_{2t}] + d_t \\ &= \tilde{E}[\mathbf{1}_{A_t} Z_t \mid \mathcal{H}_{2t}] + d_t \\ &\leq \tilde{E}[\mathbf{1}_D Z_t \mid \mathcal{H}_{2t}] + \tilde{E}[d_t \mid \mathcal{H}_{2t}] + d_t, \end{aligned}$$

where the first and third lines use $Z_t \leq 1$ and the second uses the measurability of A_t with respect to $\mathcal{H}_{2t}$. All the terms on the last line converge $\tilde{P}$-almost surely to zero, and so $\tilde{E}[Z_t|\mathcal{H}_{2t}] \to 0$ $\tilde{P}$-a.s. on the set D. Egorov's Theorem (Chung (1974, p. 74)) then implies that there exists $F \subset D$ such that $\tilde{P}(F) > 0$ on which the convergence of p_t and $\tilde{E}[Z_t|\mathcal{H}_{2t}]$ is uniform.

To clarify the remainder of the argument, we present here the case of short-lived player 2 (long-lived player 2 is discussed in Appendix A.2). This case is particularly simple, because if player 2 believed his opponent was "almost" the commitment type, then in each period 2 plays the same equilibrium action as if he was *certain* he was facing the commitment simple type.

¿From the upper semi-continuity of the best response correspondence, there exists $\psi > 0$ such that for any history h_{1s} and any $\zeta_1 \in \Delta^I$ satisfying $\|\zeta_1 - \varsigma_1\| \leq \psi$, a best response to ζ_1 is also a best response to ς_1, and so necessarily equals ς_2. The uniform convergence of $\tilde{E}[Z_t|\mathcal{H}_{2t}]$ on F implies that, for any $\xi > 0$, there exists a time T such that on F, for all $t > T$, $p_t > \eta$ and (since $\hat{\sigma}_{1s} = \varsigma_1$)

$$\tilde{E}\left[\sup_{s\geq t}\left\|\varsigma_1 - \tilde{E}[\tilde{\sigma}_{1s}|\mathcal{H}_{2s}]\right\|\,\middle|\,\mathcal{H}_{2t}\right] < \xi\psi.$$

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, implying (8). ∎

4.3 Player 1's Beliefs about Player 2's Future Behavior

Our next step is to show that, with positive probability, player 1 eventually expects player 2 to play a best response to the commitment type for the remainder of the game. We first show that, while player 2's private history h_{2t} is typically of use to player 1 in predicting 2's period-s behavior for $s > t$, this usefullness vanishes as $s \to \infty$. The intuition is straightforward. If period-s behavior is eventually (as s becomes large) independent of h_{2t}, then clearly h_{2t} is eventually of no use in predicting that behavior. Suppose then that h_{2t} is essential to predicting player 2's behavior in all periods $s > t$. Then, player 1 continues to receive information about this history from subsequent observations, reducing the value of having h_{2t} explicitly revealed. As time passes player 1 will figure out whether h_{2t} actually occurred from her own observations, again reducing the value of independently knowing h_{2t}.

Denote by $\beta(\mathcal{A}, \mathcal{B})$ the smallest σ-algebra containing 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.

Lemma 3 *Suppose Assumptions 1 and 3 hold. For any $t > 0$ and $\tau \geq 0$,*

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

Proof. We prove the result here for $\tau = 0$. The case of $\tau \geq 1$ is proved by induction in Appendix A.3. Suppose $K \subset J^t$ is a set of t-period player 2 action profiles $(j_0, j_1, ..., j_{t-1})$. We also denote by K the corresponding event (i.e., subset of Ω). 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

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

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,

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

Since the sequence of random variables $\{\tilde{P}[K|\mathcal{H}_{1s}]\}_s$ is a martingale relative to $(\{\mathcal{H}_{1s}\}_s, \tilde{P})$, it converges $\tilde{P}$-almost surely to a non-negative limit $\tilde{P}\left[K|\mathcal{H}_{1\infty}\right]$ as $s\to\infty$. Consequently, the left side of this inequality converges $\tilde{P}$-almost surely to zero. The signals generated by player 2's actions satisfy Assumption 3, so an identical argument to that given at the end of the proof of Lemma 1 in Cripps, Mailath, and Samuelson (2004) establishes that $\tilde{P}$-almost everywhere on K,

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

where $\beta\left(\mathcal{A},B\right)$ is the smallest σ-algebra containing both the σ-algebra $\mathcal{A}$ and the event B. Moreover, $\tilde{P}\left[K|\mathcal{H}_{1\infty}\right](\omega)>0$ for $\tilde{P}$-almost all $\omega\in K$. Thus, $\tilde{P}$-almost everywhere on K,

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

Since this holds for all $K\in\mathcal{H}_{2t}$,

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

giving the result for $\tau=0$. ■

Now we apply Lemma 3 to a particular piece of information player 2 could have at time t. By Lemma 2, with positive probability, we reach a time t at which player 2 assigns high probability to the event that all his future behavior is a best reply to the commitment type. Intuitively, by Lemma 3, these period-t beliefs of player 2 about his own future behavior will, eventually, become known to player 1.

This step is motivated by the observation that, if player 1 eventually expects player 2 to always play a best response to the commitment type, then the normal type of player 1 will choose to deviate from the behavior of the commitment type (which is not a best response to player 2's best response to the commitment type). At this point, we appear to have a contradiction between player 2's belief on the event F (from Lemma 2) that the normal and commitment types are playing identically and player 1's behavior on the event $F^\dagger$ (the event where player 1 expects player 2 to always play a best response to the commitment type, identified in the next lemma). This contradiction would be immediate if $F^\dagger$ was both a subset of F and measurable for player 2. Unfortunately we have no reason to expect either. However, the next lemma shows that $F^\dagger$ is in fact close to a $\mathcal{H}_{2s}$-measurable set on which player 2's beliefs that player 1 is the commitment type do not converge to zero. In this case we will (eventually) have a contradiction: On all such histories, the normal and commitment types are playing identically. However, nearly everywhere on a relatively large subset of these states, player 1 is deviating from the commitment strategy in an identifiable way.

Recall that $j^*(h_s)$ is the action played for sure in period s after the public history h_s by player 2's best response to the commitment type. Hence, $\tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\mathcal{H}_{1s}]$ is the probability player 1 assigns in period s to the event that 2 best responds to the commitment type in period $s' \geq s$. For the case of the short-lived uninformed players and the simple commitment type, $j^*(h_s) = \varsigma_2$ for all h_s,[13] and so $\left\|\tilde{E}[\sigma_{2s'}|\mathcal{H}_{1s}] - \varsigma_2\right\| \geq 1 - \tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\mathcal{H}_{1s}]$. So, in that case, (12) implies $\left\|\tilde{E}[\sigma_{2s'}|\mathcal{H}_{1s}] - \varsigma_2\right\| < \nu$.

Lemma 4 *Suppose the hypotheses of Theorem 2 hold, and suppose there is a Nash equilibrium in which reputations do not necessarily disappear, i.e.,* $\tilde{P}(\{p_t \not\to 0\}) > 0$. *Let* $\eta > 0$ *be the constant and* F *the positive probability event identified in Lemma 2. For any* $\nu > 0$ *and number of periods* $\tau > 0$, *there exists an event* $F^\dagger$ *and a time* $T(\nu, \tau)$ *such that for all* $s > T(\nu, \tau)$ *there exists* $C_s^\dagger \in \mathcal{H}_{2s}$ *with:*

$$p_s > \eta \text{ on } C_s^\dagger, \tag{9}$$

$$F^\dagger \cup F \subset C_s^\dagger, \tag{10}$$

$$\tilde{P}(F^\dagger) > \tilde{P}(C_s^\dagger) - \nu \tilde{P}(F), \tag{11}$$

[13] Here we use ς_2 to denote the pure action receiving probability one under ς_2.

and for any $s' \in \{s, s+1, ..., s+\tau\}$*, on* $F^\dagger$*,*

$$\tilde{E}[\ \sigma_{2s'}^{j^*(h_{s'})} \mid \mathcal{H}_{1s}\ ] > 1 - \nu, \qquad \tilde{P}\text{-a.s.} \tag{12}$$

Proof. Fix $\nu \in (0,1)$ and a number of periods $\tau > 0$. Fix $\xi < (\frac{1}{4}\nu\tilde{P}(F))^2$, and let T denote the critical period identified in Lemma 2 for this value of ξ.

Player 1's minimum estimated probability on $j^*(h_{s'})$ over periods $s, \ldots, s+\tau$ can be written as $f_s \equiv \min_{s \le s' \le s+\tau} \tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\mathcal{H}_{1s}]$. Notice that $f_s > 1 - \nu$ is a sufficient condition for inequality (12).

The first part of the proof is to find a lower bound for f_s. For any $t \le s$, the triangle inequality implies

$$1 \ge f_s \ge \min_{s \le s' \le s+\tau} \tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - k_s^t,$$

where $k_s^t \equiv \max_{s \le s' \le s+\tau} |\tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})] - \tilde{E}[\sigma_{2s'}^{j^*(h_{s'})}|\mathcal{H}_{1s}]|$ for $t \le s$. By Lemma 3, $\lim_{s\to\infty} k_s^t = 0$ $\tilde{P}$-almost surely.

As $\sigma_{2s'}^{j^*(h_{s'})} \le 1$ and is equal to 1 on G_t^o, the above implies

$$f_s \ge \tilde{P}\left(G_t^o \mid \beta(\mathcal{H}_{1s}, \mathcal{H}_{2t})\right) - k_s^t.$$

Moreover, the sequence of random variables $\{\tilde{P}(G_t^o|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t}))\}_s$ is a martingale with respect to the filtration $\{\mathcal{H}_{1s}\}_s$, and so converges almost surely to a limit, $g^t \equiv \tilde{P}(G_t^o|\beta(\mathcal{H}_{1\infty}, \mathcal{H}_{2t}))$. Hence

$$1 \ge f_s \ge g^t - k_s^t - \ell_s^t, \tag{13}$$

where $\ell_s^t \equiv |g^t - \tilde{P}(G_t^o|\beta(\mathcal{H}_{1s}, \mathcal{H}_{2t}))|$ and $\lim_{s\to\infty} \ell_s^t = 0$ $\tilde{P}$-almost surely.

The second step of the proof determines the sets $C_s^\dagger$ and a set that we will use to later determine $F^\dagger$. For any $t \ge T$, define

$$K_t \equiv \{\omega : \tilde{P}(G_t^o \mid \mathcal{H}_{2t}) > 1 - \xi\ ,\ p_t > \eta\} \in \mathcal{H}_{2t}.$$

Let F_t^s denote the event $\cap_{\tau=t}^s K_\tau$ and set $F_t \equiv \cap_{\tau=t}^\infty K_\tau$; note that $\liminf K_t \equiv \cup_{t=T}^\infty \cap_{\tau=t}^\infty K_\tau = \cup_{t=T}^\infty F_t$. By Lemma 2, $F \subset K_t$ for all $t \ge T$, so $F \subset F_t^s$, $F \subset F_t$, and $F \subset \liminf K_t$.

Define $N_t \equiv \{\omega : g^t \ge 1 - \sqrt{\xi}\}$. Set $C_s^\dagger \equiv F_T^s \in \mathcal{H}_{2s}$ and define an intermediate set F^* by $F^* \equiv F_T \cap N_T$. Because $C_s^\dagger \subset K_s$, (9) holds. In addition, $F^* \cup F \subset C_s^\dagger$, and hence (10) holds with F^* in the role of $F^\dagger$. By definition,

$$\tilde{P}(C_s^\dagger) - \tilde{P}(F^*) = \tilde{P}(C_s^\dagger \cap \overline{(F_T \cap N_T)}) = \tilde{P}((C_s^\dagger \cap \bar{F}_T) \cup (C_s^\dagger \cap \bar{N}_T)),$$

where we use bars to denote complements. By our choice of $C_s^\dagger$, the event $C_s^\dagger \cap \bar{N}_T$ is a subset of the event $K_T \cap \bar{N}_T$. Thus, we have the bound

$$\tilde{P}(C_s^\dagger) - \tilde{P}(F^*) \leq \tilde{P}(C_s^\dagger \cap \bar{F}_T) + \tilde{P}(K_T \cap \bar{N}_T). \tag{14}$$

We now find upper bounds for the two terms on the right side of (14). First notice that $\tilde{P}(C_s^\dagger \cap \bar{F}_T) = \tilde{P}(F_T^s) - \tilde{P}(F_T)$. Since $\lim_{s\to\infty} \tilde{P}(F_T^s) = \tilde{P}(F_T)$, there exists $T' \geq T$ such that

$$\tilde{P}(C_s^\dagger \cap \bar{F}_T) < \sqrt{\xi} \quad \text{for all} \quad s \geq T'. \tag{15}$$

Also, as $\tilde{P}(G_t^o|K_t) > 1 - \xi$ and $K_t \in \mathcal{H}_{2t}$, the properties of iterated expectations imply that $1 - \xi < \tilde{P}(G_t^o|K_t) = \tilde{E}[g^t|K_t]$. Since $g^t \leq 1$, we have

$$\begin{aligned} 1 - \xi < \tilde{E}[g^t \mid K_t] &\leq (1 - \sqrt{\xi})\tilde{P}(\bar{N}_t \mid K_t) + \tilde{P}(N_t \mid K_t) \\ &= 1 - \sqrt{\xi}\tilde{P}(\bar{N}_t \mid K_t). \end{aligned}$$

The extremes of the above inequality imply that $\tilde{P}(\bar{N}_t|K_t) < \sqrt{\xi}$. Hence, taking $t = T$ we get

$$\tilde{P}(K_T \cap \bar{N}_T) < \sqrt{\xi}. \tag{16}$$

Using (15) and (16) in (14), $\tilde{P}(C_s^\dagger) - \tilde{P}(F^*) < 2\sqrt{\xi}$ for all $s \geq T'$. Given $F \subset C_s^\dagger$, the bound on ξ, and $\nu < 1$, it follows that

$$\tilde{P}(F^*) > \tilde{P}(F) - 2\sqrt{\xi} > \frac{1}{2}\tilde{P}(F) > 0.$$

Finally, we combine the two steps above to obtain $F^\dagger$. As $\tilde{P}(F^*) > 0$ and $k_s^T + \ell_s^T$ converges almost surely to zero, by Egorov's Theorem, there exists $F^\dagger \subset F^*$ such that $\tilde{P}(F^* \setminus F^\dagger) < \sqrt{\xi}$ and a time $T'' > T$ such that $k_s^{T'} + \ell_s^{T'} < \sqrt{\xi}$ on $F^\dagger$ for all $s \geq T''$. Since $F^\dagger \cup F \subset F^* \cup F \subset C_s^\dagger$, (10) holds. Let $T(\nu, \tau) \equiv \max\{T'', T'\}$. Also, $g^T \geq 1 - \sqrt{\xi}$ on $F^\dagger$, because $F^\dagger \subset N_T$. Hence on $F^\dagger$, by (13), $f_s > 1 - 2\sqrt{\xi}$ for all $s > T(\nu, \tau)$. This, and the bound on ξ, implies (12). Moreover, as $\tilde{P}(F^* \setminus F^\dagger) < \sqrt{\xi}$ and $\tilde{P}(C_s^\dagger) - \tilde{P}(F^*) < 2\sqrt{\xi}$, (11) holds for all $s > T(\nu, \tau)$. ■

When player 2 is long-lived, it will be convenient to know that the conclusions of Lemma 4 hold on a sequence of cylinder sets:

Corollary 2 *Assume the conditions of Lemma 4. Define* $F_s^\dagger = \{\omega \in \Omega : \mathrm{proj}_s(\omega) = \mathrm{proj}_s(\omega')$ *for some* $\omega' \in F^\dagger\}$, *where* $\mathrm{proj}_s(\omega)$ *is the projection of* ω *onto* $(I \times J \times Y)^s$. *Then, (10), (11), and (12) hold for* $F_s^\dagger$ *replacing* $F^\dagger$.

Proof. The proof follows from the observation that, for all s, $F^\dagger \subset F_s^\dagger \subset C_s^\dagger$ (since $C_s^\dagger \in \mathcal{H}_{2s}$) and (12) is a condition that is $\mathcal{H}_{1s}$-measurable. ■

4.4 Toward a Contradiction

We have shown that when reputations do not necessarily disappear, there exists a set $F^\dagger$ on which (12) holds and $F^\dagger \subset C_s^\dagger \in \mathcal{H}_{2s}$. The remaining argument is more transparent in the setting of the short-lived player 2s of Theorem 1. Accordingly, we first prove Theorem 1, and then give the distinct argument needed when player 2 is long-lived and the commitment strategy is not simple.

In broad brushstrokes, the argument proving Theorem 1 is as follows. First, we conclude that on $F^\dagger$, the normal type will not be playing the commitment strategy. To be precise—on $F^\dagger$ there will exist a stage-game action played by ς_1 but not by the normal type. This will bias player 2's expectation of the normal type's actions away from the commitment strategy on $C_s^\dagger$, because there is little probability weight on $C_s^\dagger \setminus F^\dagger$. We then get a contradiction, because the fact that $p_s > \eta$ on $C_s^\dagger$ implies player 2 must believe the commitment type's strategy and the normal type's average strategy are the same on $C_s^\dagger$.

The argument proving Theorem 2 must deal with the nonstationary nature of the commitment strategy (and the nonstationary nature of the failure of credibility). As in the simple case, we have found a set of states $F^\dagger$ 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 to this is not the commitment strategy, and hence the normal type does not play the commitment strategy. We will derive a contradiction by showing that player 2 almost comes to know this.

The complication is that it may be very difficult for player 2 to predict just how the normal type's strategy deviates from the commitment strategy. When working with the stationary commitment strategy of Theorem 1, we can be certain there is a stage-game action played by the commitment type which the normal type's strategy would (eventually) not play after any private history. In the setting of Theorem 2, however, the normal type's deviation from the nonstationary commitment strategy may be much more complicated, and may depend on private (rather than just public) information.

4.5 Proof of Theorem 1

Suppose, en route to the contradiction, that there is a Nash equilibrium in which reputations do not necessarily disappear. Then $\tilde{P}(\{p_t \nrightarrow 0\}) > 0$. Let $\underline{\varsigma}_1 \equiv \min_{i \in I}\{\varsigma_1^i : \varsigma_1^i > 0\}$, that is, $\underline{\varsigma}_1$ is the smallest non-zero probability attached to an action under the commitment strategy ς_1. Since $(\varsigma_1, \varsigma_2)$ is not a Nash equilibrium,

ς_1 plays an action that is suboptimal by at least $\gamma > 0$ when player 2 uses any strategy sufficiently close to ς_2. That is, there exists $\gamma > 0$, $i' \in I$ with $\varsigma_1^{i'} > 0$ and $\bar{\nu} > 0$ such that

$$\gamma < \min_{\|\sigma_2 - \varsigma_2\| \leq \bar{\nu}} \left(\max_{i \in I} \pi_1(i, \sigma_2) - \pi_1(i', \sigma_2) \right).$$

Finally, for a given discount factor $\delta_1 < 1$ there exists a τ sufficiently large such that the loss of γ for one period is larger than any feasible potential gain deferred by τ periods: $(1 - \delta_1)\gamma > \delta_1^\tau 2 \max_{ij} |\pi_1(i, j)|$.

Fix the event F from Lemma 2. For $\nu < \min\{\bar{\nu}, \frac{1}{2}\underline{\varsigma}_1\}$ and τ above, let $F^\dagger$ and, for $s > T(\nu, \tau)$, $C_s^\dagger$ be the events described in Lemma 4. Now consider the normal type of player 1 in period $s > T(\nu, \tau)$ at some state in $F^\dagger$. By (12), she expects player 2 to play within $\nu < \bar{\nu}$ of ς_2 for the next τ periods. Playing the action i' is conditionally dominated in period s, since the most she can get from playing i' in period s is worse than playing a best response to ς_2 for τ periods and then being minmaxed. Thus, on $F^\dagger$ the normal type plays action i' with probability zero: $\sigma_{1s}^{i'} = 0$.

Now we calculate a lower bound on the difference between player 2's beliefs about the normal type's probability of playing action i' in period s, $\tilde{E}[\sigma_{1s}^{i'}|\mathcal{H}_{2s}]$, and the probability the commitment type plays action i' on the set of states $C_s^\dagger$:

$$\begin{aligned}
\tilde{E}\left[\left|\varsigma_1^{i'} - \tilde{E}[\sigma_{1s}^{i'}|\mathcal{H}_{2s}]\right| \mathbf{1}_{C_s^\dagger}\right] &\geq \tilde{E}\left[\left(\varsigma_1^{i'} - \tilde{E}[\sigma_{1s}^{i'}|\mathcal{H}_{2s}]\right) \mathbf{1}_{C_s^\dagger}\right] \\
&\geq \underline{\varsigma}_1 \tilde{P}(C_s^\dagger) - \tilde{E}\left[\sigma_{1s}^{i'} \mathbf{1}_{C_s^\dagger}\right] \\
&\geq \underline{\varsigma}_1 \tilde{P}(C_s^\dagger) - \left(\tilde{P}(C_s^\dagger) - \tilde{P}(F^\dagger)\right) \\
&\geq \underline{\varsigma}_1 \tilde{P}(C_s^\dagger) - \nu \tilde{P}(F) \\
&\geq \frac{1}{2}\underline{\varsigma}_1 \tilde{P}(F). \qquad (17)
\end{aligned}$$

The first inequality above follows from removing the absolute values. The second inequality applies $\varsigma_1^{i'} \geq \underline{\varsigma}_1$, uses the $\mathcal{H}_{2s}$-measurability of $C_s^\dagger$ and applies the properties of conditional expectations. The third applies the fact that $\sigma_{1s}^{i'} = 0$ on $F^\dagger$ and $\sigma_{1s}^{i'} \leq 1$. The fourth inequality applies (11) in Lemma 4. The fifth inequality follows $\nu < \frac{1}{2}\underline{\varsigma}_1$ and $F \subset C_s^\dagger$ (by (10)).

¿From Corollary 1, $p_s \|\varsigma_1 - \tilde{E}(\tilde{\sigma}_{1s}|\mathcal{H}_{2s})\| \to 0$ $\tilde{P}$-almost surely. It follows that

$$p_s |\varsigma_1^{i'} - \tilde{E}(\tilde{\sigma}_{1s}^{i'}|\mathcal{H}_{2s})| \mathbf{1}_{C_s^\dagger} \to 0, \qquad \tilde{P}\text{–a.s.}$$

But, by Lemma 4, $p_s > \eta$ on the set $C_s^\dagger$, and so

$$|\varsigma_1^{i'} - \tilde{E}(\tilde{\sigma}_{1s}^{i'}|\mathcal{H}_{2s})|\mathbf{1}_{C_s^\dagger} \to 0, \qquad \tilde{P}\text{−a.s.}$$

This concludes the proof of Theorem 1, since we now have a contradiction with $\tilde{P}(F) > 0$ (from Lemma 2) and (17), which holds for all $s > T(\nu, \tau)$.

4.6 Proof of Theorem 2

We first argue that, after any sufficiently long public history, there is one continuation public history after which the commitment type will play some action $i^o \in I$ with positive probability, but after which the normal type will choose not to play i^o, regardless of her private history. To find such a history, notice that $\hat{\sigma}_2$ (player 2's best response to the commitment strategy) is pure and therefore public, ensuring that the normal player 1 has a public best response to $\hat{\sigma}_2$ and that it is not $\hat{\sigma}_1$. Hence, there exists a public history where 1's public best response differs from the commitment strategy, for all private histories consistent with this public history. If we can show this preference is strict, this will still hold when player 2 is just playing close to a best response, which will open the door to a contradiction. The formal statement is (the proof is in Appendix A.4):

Lemma 5 *Suppose $\hat{\sigma}_1$ is a public strategy with no long-run credibility (with an associated T^o), and $\hat{\sigma}_2$ is player 2's public best reply. Then, player 1 has a public best reply, $\sigma_1^\dagger$, to $\hat{\sigma}_2$. There exists $\hat{\tau} \in \mathbb{N}$, $\lambda > 0$, and $\kappa > 0$ such that for all $s > T^o$ and each $h_s \in H_s$, there is an action i^o, a period $s' \leq s + \hat{\tau}$, and a public continuation history $h_{s'}^o$ of h_s, such that*

1. *$\hat{\sigma}_{1s'}^{i^o}(h_{s'}^o) \geq \lambda$,*
2. *the action i^o receives zero probability under $\sigma_1^\dagger(h_{s'}^o)$, and*
3. *player 1's payoff from playing i^o and continuing with strategy $\hat{\sigma}_1$ is at least κ less that what she gets from playing $\sigma_1^\dagger$ at $h_{s'}^o$, i.e.,*

$$E^{(\sigma_1^\dagger,\hat{\sigma}_2)}[\pi_{1s'}|h_{s'}^o] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[(1-\delta_1)\pi_1(i^o, j_{s'}) + \delta_1\pi_{1,s'+1}|h_{s'}^o\right] \geq \kappa.$$

For $s > T^o$, Lemma 5 describes how player 1's best response to $\hat{\sigma}_2$ differs from $\hat{\sigma}_1$. In the game with incomplete information, Lemma 5 defines three $\mathcal{H}_s$-measurable functions, $i(\cdot; s) : \Omega \to I$, $s'(\cdot; s) : \Omega \to \{t : s \leq t \leq s + \tau\}$, and $\mathfrak{h}(\cdot;s) : \Omega \to \cup_{t=0}^{\infty} Y^t$ as follows: Associated with each state $\omega \in \Omega$ is the implied s-period public history, h_s. The action-period pair $(i(\omega;s), s'(\omega;s))$ is the action-period pair (i^o, s') from Lemma 5 for the public history h_s. Finally, $\mathfrak{h}(\omega;s)$ is the $s'(\omega;s)$-period continuation history $h^o_{s'}$ of h_s from Lemma 5. We emphasize that $\mathfrak{h}(\omega;s)$ is typically *not* the $s'(\omega;s)$-period public history of ω (for a start, it is $\mathcal{H}_s$-measurable); while the first s-periods of $\mathfrak{h}(\omega;s)$ are the s-period public history of ω, the next $s'(\omega;s) - s$ periods describe the public signals from Lemma 5.

With these functions in hand, we can describe how player 1's behavior differs from that of the commitment type when she is sufficiently confident that player 2 is best responding to the commitment type (where $\underline{\rho} \equiv \min_{y,i,j} \rho^y_{ij} > 0$ and λ is from Lemma 5; the proof is in Appendix A.5):

Lemma 6 *Suppose the hypotheses of Theorem 2 hold, and suppose there is a Nash equilibrium in which reputations do not necessarily disappear, i.e.,* $\tilde{P}(\{p_t \not\to 0\}) > 0$. *Let* $\hat{\tau}$, λ, *and* κ *be the constants identified in Lemma 5, and* $M \equiv \max_{i\in I, j\in J, \ell\in\{1,2\}} |\pi_\ell(i,j)|$. *Suppose* $\tau > \hat{\tau}$ *satisfies* $12M\delta_1^\tau < \kappa$, $\nu > 0$ *satisfies* $12M\nu < \kappa\underline{\rho}^\tau$, *and* $\{F^\dagger_s\}_s$ *is the sequence of events identified in Corollary 2. For all* $s \geq T(\nu, 2\tau)$,

1. $\hat{\sigma}^{i(\omega;s)}_{1,s'(\omega;s)}(\mathfrak{h}(\omega;s)) \geq \lambda$,
2. *the set* $F^\ddagger_s \equiv \{\omega \in F^\dagger_s : h_{s'(\omega;s)}(\omega) = \mathfrak{h}(\omega;s)\}$ *has probability* $\tilde{P}(F^\ddagger_s) \geq \underline{\rho}^\tau \tilde{P}(F^\dagger_s) > 0$, *and*
3. *for all* $\omega \in F^\ddagger_s$,
$$\tilde{\sigma}^{i(\omega;s)}_{1,s'(\omega;s)}(h_{1,s'(\omega;s)}(\omega)) = 0.$$

If the events $F^\ddagger_s$ were known to player 2 in period s, then the argument is now complete, since there would be a contradiction between player 2's belief that the normal and commitment type play the same way on $F^\ddagger_s$ and player 1's actual behavior. However, $F^\ddagger_s$ is not known to player 2. On the other hand, $F^\ddagger_s$ is approximated by $C^\ddagger_s$ (the analogous modification of $C^\dagger_s$, defined below), an event known by player 2 in period s. At the same time, we must still deal with the random nature of $i(\cdot;s)$ and $s'(\cdot;\omega)$.

To complete the argument then, suppose the assumptions of Lemma 6 (including the bounds on τ and ν) hold, and in addition

$$\nu < \frac{2\lambda\underline{\rho}^\tau}{2\lambda\underline{\rho}^\tau + 3}. \tag{18}$$

The set of states consistent with 2's information at time s, $C_s^\dagger$, and the "right" continuation public history, is $C_s^\ddagger \equiv \{\omega \in C_s^\dagger : h_{s'(\omega)}(\omega) = \mathfrak{h}(\omega; s)\}$. Note that $\tilde{P}(C_s^\ddagger \backslash F_s^\ddagger) \leq \tilde{P}(C_s^\dagger \backslash F_s^\dagger)$, and since $C_s^\dagger \supset F_s^\dagger$, $C_s^\ddagger \supset F_s^\ddagger$. We also partition $C_s^\ddagger$ into the subevents corresponding to the relevant period in which the action $i = i(\omega; s)$ is not optimal: $C_s^{\ddagger it} \equiv \{\omega \in C_s^\dagger : i(\omega; s) = i,\ s'(\omega; s) = t, h_t(\omega) = \mathfrak{h}(\omega; s)\}$, so that $C_s^\ddagger = \cup_{t=s}^{s+\tau} \cup_{i \in I} C_s^{\ddagger it}$. Note that $C_s^{\ddagger it} \in \mathcal{H}_{2t}$ for all $i \in I$ and $t = s, \ldots, s+\tau$.

For each ω, let $i^o = i(\omega; s)$ and $s^o = s'(\omega; s)$. Now, for fixed ω and implied fixed action i^o and period s^o, define $\hat{f}_s(\omega) \equiv \hat{\sigma}_{1s^o}^{i^o}(\omega)$ and $\tilde{f}_s(\omega) \equiv \tilde{E}\left[\tilde{\sigma}_{1s^o}^{i^o} | \mathcal{H}_{2s^o}\right](\omega)$. In the last expression, for fixed action i^o and period s^o, $\tilde{E}\left[\tilde{\sigma}_{1s^o}^{i^o} | \mathcal{H}_{2s^o}\right]$ is the conditional expected value of $\tilde{\sigma}_{1s^o}^{i^o}$. In particular, for $\omega \in C_s^{\ddagger it}$, $s^o = t$ and $i^o = i$, and we can write $\hat{f}_s(\omega) \equiv \hat{\sigma}_{1t}^{i}(\omega)$ and $\tilde{f}_s(\omega) \equiv \tilde{E}\left[\tilde{\sigma}_{1t}^{i} | \mathcal{H}_{2t}\right](\omega)$. Then, $Z_s(\omega) \equiv \sup_{t \geq s} \left\| \hat{\sigma}_{1t} - \tilde{E}\left[\tilde{\sigma}_{1t} | \mathcal{H}_{2t}\right] \right\| \geq \left| \hat{f}_s(\omega) - \tilde{f}_s(\omega) \right|$.

So,

$$\begin{aligned}
& \tilde{E}[Z_s \mathbf{1}_{C_s^\ddagger}] \geq \tilde{E}\left[\left(\hat{f}_s - \tilde{f}_s\right) \times \mathbf{1}_{C_s^\ddagger}\right] && (19) \\
= \ & \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{E}\left[\left(\hat{f}_s - \tilde{f}_s\right) \times \mathbf{1}_{C_s^{\ddagger it}}\right] \\
= \ & \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{E}\left[\left(\hat{\sigma}_{1t}^i - \tilde{E}\left[\tilde{\sigma}_{1t}^i | \mathcal{H}_{2t}\right]\right) \times \mathbf{1}_{C_s^{\ddagger it}}\right] \\
= \ & \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{E}\left[\tilde{E}\left[\left(\hat{\sigma}_{1t}^i - \tilde{\sigma}_{1t}^i\right) \mathbf{1}_{C_s^{\ddagger it}} | \mathcal{H}_{2t}\right]\right], && (20)
\end{aligned}$$

where the last equality follows from $C_s^{\ddagger t} \in \mathcal{H}_{2t}$. Now, define $F_s^{\ddagger it} \equiv \{\omega \in F_s^\dagger : i(\omega; s) = i,\ s'(\omega; s) = t, h_t(\omega) = \mathfrak{h}(\omega; s)\}$, and so $F_s^\ddagger = \cup_{t=s}^{s+\tau} \cup_{i \in I} F_s^{\ddagger it}$. Since $F_s^\dagger \subset C_s^\dagger$, $F_s^{\ddagger it} \subset C_s^{\ddagger it}$, and so (20) is at least as large as

$$\begin{aligned}
& \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{E}\left[\tilde{E}\left[\left(\hat{\sigma}_{1t}^i - \tilde{\sigma}_{1t}^i\right) \mathbf{1}_{F_s^{\ddagger it}} | \mathcal{H}_{2t}\right]\right] - \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{P}\left(C_s^{\ddagger it} \backslash F_s^{\ddagger it}\right) \\
= \ & \tilde{E}\left[\left(\hat{f}_s \mathbf{1}_{F_s^\ddagger} - \sum_{t=s}^{s+\tau} \sum_{i \in I} \tilde{E}\left[\tilde{\sigma}_{1t}^i \mathbf{1}_{F_s^{\ddagger it}} | \mathcal{H}_{2t}\right]\right)\right] - \tilde{P}\left(C_s^\ddagger \backslash F_s^\ddagger\right) \\
= \ & \tilde{E}\left[\hat{f}_s \mathbf{1}_{F_s^\ddagger}\right] - \tilde{P}\left(C_s^\ddagger \backslash F_s^\ddagger\right) \\
> \ & \lambda \tilde{P}(F_s^\ddagger) - \tilde{P}(C_s^\dagger \backslash F_s^\dagger), && (21)
\end{aligned}$$

where the last equality is an implication of $\tilde{E}\left[\tilde{\sigma}^i_{1t}\mathbf{1}_{F_s^{\ddagger it}}|\mathcal{H}_{2t}\right] = 0$ $\tilde{P}$-almost surely. Hence, from the chain from (19) to (21), we have

$$\tilde{E}[Z_s\mathbf{1}_{C_s^\ddagger}] > \lambda\underline{\rho}^\tau\tilde{P}(F_s^\dagger) - (\tilde{P}(C_s^\dagger) - \tilde{P}(F_s^\dagger)). \tag{22}$$

Applying the bounds $\nu\tilde{P}(F) > \tilde{P}(C_s^\dagger) - \tilde{P}(F_s^\dagger)$ and $\tilde{P}(F_s^\dagger) > \tilde{P}(F)(1-\nu)$ from Corollary 2 to the right side of (22) gives

$$\tilde{E}[Z_s\mathbf{1}_{C_s^\ddagger}] > (\lambda\underline{\rho}^\tau(1-\nu) - \nu)\tilde{P}(F).$$

The bound (18) ensures that $\lambda\underline{\rho}^\tau(1-\nu) - \nu > \nu/2$, and hence

$$\tilde{E}[Z_s\mathbf{1}_{C_s^\ddagger}] > \frac{1}{2}\nu\tilde{P}(F).$$

However, $\tilde{P}(C_s^\ddagger) > \underline{\rho}^\tau(1-\nu)\tilde{P}(F) > 0$ and since $C_s^\ddagger \subset \{\omega : p_t \nrightarrow 0\}$, $Z_s\mathbf{1}_{C_s^\ddagger} \to 0$ $\tilde{P}$-almost surely, the desired contradiction.

5 Imperfect Private Monitoring

In this section, we briefly sketch how our results can be extended to the case of private monitoring. Instead of observing a public signal y at the end of each period, player 1 observes a *private* signal θ (drawn from a finite set Θ) and player 2 observes a private signal ζ (drawn from a finite set Z). A history for a player is the sequence of his or her actions and private signals. Given the underlying action profile (i, j), we let ρ_{ij} denote a probability distribution over $\Theta \times Z$. We use $\rho_{ij}^{\theta\zeta}$ to denote the probability of the signal profile (θ, ζ) conditional on (i, j). The marginal distributions are $\rho_{ij}^\theta = \sum_\zeta \rho_{ij}^{\theta\zeta}$ and $\rho_{ij}^\zeta = \sum_\theta \rho_{ij}^{\theta\zeta}$. The case of public monitoring is a special case: take $\Theta = Z$ and $\Sigma_{\theta\in\Theta}\rho_{ij}^{\theta\theta} = 1$ for all i, j.

We now describe the analogs of our earlier assumptions on the monitoring technology. The full-support assumption is:

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

Note that we do *not* assume that $\rho_{ij}^{\theta\zeta} > 0$ for all $(i, j) \in I \times J$ and $(\theta, \zeta) \in \Theta^2$ (which would rule out public monitoring). Instead, the full-support assumption is that each signal is observed with positive probability under every action profile.

Assumption 6 (IDENTIFICATION 1) *For all $j \in J$, the I columns in the matrix $(\rho_{ij}^{\zeta})_{\zeta \in Z, i \in I}$ are linearly independent.*

Assumption 7 (IDENTIFICATION 2) *For all $i \in I$, the J columns in the matrix $(\rho_{ij}^{\theta})_{\theta \in \Theta, j \in J}$ are linearly independent.*

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. This is established in Theorem 6, which is a minor extension of Fudenberg and Levine (1992).[14]

Theorem 6 *Suppose the game has imperfect private monitoring satisfying Assumptions 5 and 6. Suppose the commitment type plays the pure action i^* in every period. For all $p_0 > 0$ and all $\varepsilon > 0$, there exists $\bar{\delta} < 1$ such that for all $\delta_1 > \bar{\delta}$, player 1's expected average discounted payoff in any Nash equilibrium is at least*

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

where

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

The proof of the following extension of Theorem 1 to the private monitoring case is essentially identical to that of Theorem 1 apart from the added notational inconvenience of private signals.

Theorem 7 *Suppose the imperfect private monitoring satisfies Assumptions 5, 6, and 7 and ς_1 satisfies Assumption 4. Then at any Nash equilibrium, $p_t \to 0$ $\tilde{P}$-almost surely.*

[14] While Fudenberg and Levine (1992) explicitly assume public monitoring, under Assumption 6, their analysis also covers imperfect private monitoring. This includes games where player 1 observes no informative signal. In such a case, when there is complete information, the one-period-memory strategies that we describe as equilibria in Section 2 of Cripps, Mailath, and Samuelson (2004) are also equilibria of the game with private monitoring. We thank Juuso Välimäki for showing us how to construct such equilibria.

A Appendix

A.1 Proof of Theorem 5

Since $p_t \to 0$ $\tilde{P}$-almost surely, we have $p_t \to 1$ $\hat{P}$-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 $t' > T$,

$$\begin{aligned} 0 \;\leq\; & \int_{S\times[0,1]} [u_2(s_1, s_2) - u_2(s_1, \xi_2(s_2, p_t))] d(p_0 \hat{\rho}_t + (1 - p_0)\tilde{\rho}_t) \\ & \leq (1 - p_0) \int_{S\times[0,\varepsilon]} [u_2(s_1, s_2) - u_2(s_1, \xi_2(s_2, p_t))] d\tilde{\rho}_t \\ & \quad + p_0 \int_{S\times[1-\varepsilon,1]} [u_2(s_1, s_2) - u_2(s_1, \xi_2(s_2, p_t))] d\hat{\rho}_t + 2M\nu, \end{aligned}$$

where M is an upper bound on the magnitude of the stage-game payoffs and the first inequality follows from (6). 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,

$$\int_{S\times[0,\varepsilon]} [u_2(s_1, s_2) - u_2(s_1, \zeta_2(s_2, p_t))] d\tilde{\rho}_t \geq -\frac{2M\nu}{1 - p_0}. \tag{A.1}$$

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

$$\int_{S\times[0,1]} [u_2(s_1, s_2) - u_2(s_1, \xi_2(s_2, p_t))] d\tilde{\rho}_t \geq -\frac{2M\nu}{1 - p_0} - 2M\nu.$$

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

$$\int_S [u_2(s_1, s_2) - u_2(s_1, \xi_2'(s_2))] d\tilde{\mu}_t \geq -\frac{2M\nu}{1 - p_0} - 2M\nu. \tag{A.2}$$

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 (5) holds for all t', it also holds in the limit. If $\tilde{\mu}_\infty$ is not a correlated equilibrium, it must then be the case that for some $\xi_2'' : S_2 \to S_2$, there exists $\kappa > 0$ so that

$$\int_S [u_2(s_1, s_2) - u_2(s_1, \xi_2''(s_2))] d\tilde{\mu}_\infty < -\kappa < 0.$$

But then for t_k sufficiently large,

$$\int_S [u_2(s_1,s_2) - u_2(s_1,\zeta_2''(s_2))]d\tilde{\mu}_{t_k} < \frac{-\kappa}{2} < 0,$$

contradicting (A.2) for ν sufficiently small.

A.2 Completion of the Proof of Lemma 2

Turning to the general case, let $M \equiv \max_{i\in I, j\in J, \ell\in\{1,2\}} |\pi_\ell(i,j)|$, so that M is an upper bound on the magnitude of stage-game payoffs. Let $\alpha = \varepsilon^o/6M$, where ε^o is given by Definition 5. If $Z_t \leq \alpha$, player 2's expected continuation payoffs at h_{2s} under the strategy profile $(\tilde{\sigma}_1, \hat{\sigma}_1, \sigma_2)$ are within $2M\alpha$ of his continuation payoff under the profile $(\hat{\sigma}_1, \hat{\sigma}_1, \sigma_2)$. Hence, if $Z_t \leq \alpha$ and history h_{2s} (for $s \geq t \geq T^o$) occurs with positive probability, then

$$\left| E^{(\tilde{\sigma}_1,\hat{\sigma}_1,\sigma_2)}[\pi_{2s} \mid h_{2s}] - E^{(\hat{\sigma}_1,\hat{\sigma}_1,\sigma_2)}[\pi_{2s} \mid h_{2s}] \right| < 2M\alpha. \tag{A.3}$$

for all σ_2.

We now show that if $Z_t \leq \alpha$ for $t \geq T^o$, then player 2 plays the pure action $j^*(h_s)$ in all future periods. Suppose instead that the equilibrium σ_2 plays $j \neq j^*(h_s)$ with positive probability in period s under a history h_{2s}. Define σ_2' to be identical to σ_2 except that, after the history h_{2s}, it places zero probability weight on the action $j^*(h_s)$ and increases the probability of all other actions played by σ_2 by equal weight. Let $\hat{\sigma}_2$ be player 2's best response to the commitment type. Then, if $Z_t \leq \alpha$ we have[15]

$$\begin{aligned} E^{(\tilde{\sigma}_1,\hat{\sigma}_1,\sigma_2)}[\pi_{2s} \mid h_{2s}] &= E^{(\tilde{\sigma}_1,\hat{\sigma}_1,\sigma_2')}[\pi_{2s} \mid h_{2s}] \\ &\leq E^{(\hat{\sigma}_1,\hat{\sigma}_1,\sigma_2')}[\pi_{2s} \mid h_{2s}] + 2M\alpha \\ &\leq E^{(\hat{\sigma}_1,\hat{\sigma}_1,\hat{\sigma}_2)}[\pi_{2s} \mid h_{2s}] - \varepsilon^o + 2M\alpha \\ &\leq E^{(\tilde{\sigma}_1,\hat{\sigma}_1,\hat{\sigma}_2)}[\pi_{2s} \mid h_{2s}] - \varepsilon^o + 4M\alpha. \end{aligned}$$

As $4M\alpha < \varepsilon^o$, $\hat{\sigma}_2$ is a profitable deviation after the history h_{2s} for player 2—a contradiction. Hence on the event $Z_t \leq \alpha$ player 2 plays $j^*(h_s)$ in all future periods. Equivalently, we have shown $\{Z_t \leq \alpha\} \subset G_t^o$. Choose $T \geq T^o$ such that $p_t > \eta$ and $\tilde{E}[Z_t|\mathcal{H}_{2t}] < \alpha\xi$ for all $t > T$. Condition (8) now follows from $\tilde{P}[\{Z_t > \alpha\} \mid \mathcal{H}_{2t}] < \xi$ for all $t > T$ on F.

[15] The equality applies the fact that in equilibrium, player 2 is indifferent between actions played with positive probability. The first inequality applies (A.3). The second inequality applies Definition 5.1. The third inequality applies (A.3) again.

A.3 Completion of the Proof of Lemma 3

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

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

where $h_{1,z+1} = (h_{1z}, y_z, i_z)$. 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, applying the triangle inequality and rearranging, we find that the right side is larger than

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

¿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 $z \in \{s, ..., s+\tau-1\}$, 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 $z = s + \tau$.

A.4 Proof of Lemma 5

Since $\hat{\sigma}_1$ is public, player 2 has a best reply $\hat{\sigma}_2$ that is public, and so player 1 has a public best reply $\sigma_1^\dagger$ to $\hat{\sigma}_2$. By Definition 5.2, for every s-period public history h_s, $s > T^o$, we have

$$E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1s}|h_s\right] > E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s}|h_s\right] + \varepsilon^o. \tag{A.4}$$

Since $\sigma_1^\dagger$ is a best response to $\hat{\sigma}_2$, player 1's payoff $E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1s}|h_s\right]$ is unchanged if the period-s mixture $\sigma_1^\dagger(h_s)$ is replaced by any other mixture that remains within the support of $\sigma_1^\dagger(h_s)$, and thereafter play continues according to $\sigma_1^\dagger$.

For $s > T^o$ and $h_s \in H_s$, let $\Upsilon(h_s)$ be the set of public histories $h_{s'}$, $s' \geq s$, that are continuations of h_s and s' is the first period in which there is an action in I receiving positive probability under $\hat{\sigma}_1$ but receiving zero probability under $\sigma_1^\dagger$.[16] Note that $\Upsilon(h_s)$ is at most countable. In addition, there are no two elements of $\Upsilon(h_s)$ with the property that one is a continuation of the other. For $h_{s'} \in \Upsilon(h_s)$, $s' > s$, in period s, every action that receives positive probability under strategy $\hat{\sigma}_1$ also receives positive probability under $\sigma_1^\dagger$, and so the comment after equation (A.4) implies

$$E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1s}|h_s\right] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s}|h_s\right] =$$
$$\sum_{i\in I}\hat{\sigma}_1^i(h_s)\delta_1\left[E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1,s+1}|(h_s,i)\right] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1,s+1}|(h_s,i)\right]\right].$$

Applying this reasoning iteratively allows us to rewrite (A.4) as

$$\varepsilon^o < \sum_{h_{s'}\in\Upsilon(h_s)}\hat{Q}(h_{s'}|h_s)\delta_1^{s'-s}\left[E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1s'}|h_{s'}\right] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s'}|h_{s'}\right]\right] \tag{A.5}$$

where $\hat{Q}(h_{s'}|h_s)$ is the probability of $h_{s'}$ given h_s under $(\hat{\sigma}_1,\hat{\sigma}_2)$.[17]

Choose $\hat{\tau}$ such that $2M\delta_1^{\hat{\tau}} < \varepsilon^o/3$. The terms in (A.5) corresponding to histories longer than $s+\hat{\tau}$ can then collectively contribute at most $\varepsilon^o/3$ to the sum.

---

[16]Because $\sigma_1^\dagger$ is a best response to $\hat{\sigma}_2$, there must exist such histories, since otherwise every action accorded positive probability by $\hat{\sigma}_1$ would be optimal, contradicting (A.4).

[17]It is possible that $\sum_{h_{s'}\in\Upsilon(h_s)}\hat{Q}(h_{s'}|h_s) < 1$. However, expected payoffs under $(\sigma_1^\dagger,\hat{\sigma}_2)$ and $(\hat{\sigma}_1,\hat{\sigma}_2)$ are equal after any history not in $\Upsilon(h_s)$, and so such histories can then be omitted from (A.5).

The remaining terms must then sum to at least $2\varepsilon^o/3$. Letting $\Upsilon(h_s;\hat{\tau})$ denote the set of histories in $\Upsilon(h_s)$ no longer than $s+\hat{\tau}$, we have

$$\frac{2\varepsilon^o}{3} < \sum_{\Upsilon(h_s;\hat{\tau})} \hat{Q}(h_{s'}|h_s)\delta_1^{s'-s}\left[E^{(\sigma_1^\dagger,\hat{\sigma}_2)}\left[\pi_{1s'}|h_{s'}\right] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s'}|h_{s'}\right]\right].$$

Let $\Upsilon^*(h_s;\hat{\tau})$ be the histories in $\Upsilon(h_s;\hat{\tau})$ satisfying

$$E^{(\sigma_1^\dagger,\hat{\sigma}_2)}[\pi_{1s'}|h_{s'}] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s'}|h_{s'}\right] \geq \frac{\varepsilon^o}{3}. \tag{A.6}$$

Then,

$$\frac{2\varepsilon^o}{3} < \hat{Q}(\Upsilon^*(h_s;\hat{\tau})|h_s)2M + (1-\hat{Q}(\Upsilon^*(h_s;\hat{\tau})|h_s))\frac{\varepsilon^o}{3},$$

and so

$$\hat{Q}(\Upsilon^*(h_s;\hat{\tau})|h_s) > q \equiv \frac{\varepsilon^o}{6M-\varepsilon^o}$$

(the denominator is positive, since Definition 5 implies $\varepsilon^o \leq 2M$).

There are at most $Y^{\hat{\tau}}$ histories in $\Upsilon^*(h_s;\hat{\tau})$. In the last period of each such history, there is an action $i \in I$ that is played with positive probability by $\hat{\sigma}_1$ and zero probability by $\sigma_1^\dagger$. Since there are at most I such actions, there is a history $h^o_{s'}(h_s) \in \Upsilon^*(h_s;\hat{\tau})$ and action $i^o(h_s)$ such that, under $(\hat{\sigma}_1,\hat{\sigma}_2)$, the probability that the history $h^o_{s'}(h_s)$ occurs and is followed by action $i^o(h_s)$ is at least $\lambda \equiv q/(IY^{\hat{\tau}})$. Trivially, then, $\hat{\sigma}_{1s'}^{i^o}(h^o_{s'}) \geq \lambda$.

Finally, since

$$\begin{aligned} E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[\pi_{1s'}|h^o_{s'}\right] \;\leq\; & \lambda E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[(1-\delta_1)\pi_1(i^o,j_{s'}) + \delta_1\pi_{1,s'+1}|h^o_{s'}\right] \\ & + (1-\lambda)E^{(\sigma_1^\dagger,\hat{\sigma}_2)}[\pi_{1s'}|h_{s'}], \end{aligned}$$

from (A.6), we have

$$E^{(\sigma_1^\dagger,\hat{\sigma}_2)}[\pi_{1s'}|h^o_{s'}] - E^{(\hat{\sigma}_1,\hat{\sigma}_2)}\left[(1-\delta_1)\pi_1(i^o,j_{s'}) + \delta_1\pi_{1,s'+1}|h^o_{s'}\right] \geq \frac{\varepsilon^o}{3\lambda} \equiv \kappa.$$

A.5 Proof of Lemma 6

We prove only the second and third assertions (the first being an immediate implication of Lemma 5 and the definitions of i, s', and $\mathfrak{h}$).

Since $\omega \in F_s^\dagger$ and $\mathrm{proj}_s(\omega') = \mathrm{proj}_s(\omega)$ implies $\omega' \in F_s^\dagger$, for any s-period public history consistent with a state in $F_s^\dagger$, and any s'-period $(s' > s)$ public continuation of that history, there is at least one state in $F_s^\dagger$ consistent with that continuation. Consequently, since every τ period public history has probability at least $\underline{\rho}^\tau$, $\tilde{P}(F_s^\ddagger) \geq \underline{\rho}^\tau \tilde{P}(F_s^\dagger) > \underline{\rho}^\tau (1-\nu)\tilde{P}(F) > 0$.

After any public history, the normal type's payoffs under $(\sigma_1^\dagger, \hat{\sigma}_2)$ are independent of her private histories—she is playing her public best response to a public strategy. At states in $F_s^\ddagger$, from Corollary 2, under $\tilde{\sigma}_1$, player 1 expects player 2's future play (over the periods $s, s+1, ..., s+2\tau$) to be within ν of his best response to the commitment strategy, $\hat{\sigma}_2$. Hence, on $F_s^\ddagger$, player 1 expects that player 2's future play (over the periods $s, s+1, ..., s+2\tau$) to be within $\nu\underline{\rho}^{-\tau}$ of his best response to the commitment strategy, $\hat{\sigma}_2$, irrespective of her play in those periods. Discounted to the period $s' \leq s+\tau$, payoffs from periods after $s+2\tau$ can differ by at most $2M\delta_1^\tau$. Hence, for states in $F_s^\ddagger$, and for any σ_1,

$$\left| E^{(\sigma_1,\sigma_2)}[\pi_{1s'} \mid \mathcal{H}_{1s'}] - E^{(\sigma_1,\hat{\sigma}_2)}[\pi_{1s'} \mid \mathcal{H}_{1s'}] \right| \leq (\nu\underline{\rho}^{-\tau} + \delta_1^\tau)2M < \kappa/3.$$

Lemma 5.3 and the restrictions on τ and ν then imply, for $\omega \in F_s^\ddagger$,

$$E^{(\sigma_1^\dagger,\sigma_2)}[\pi_{1s'}|\mathcal{H}_{1s'}] \geq \frac{\kappa}{3} + E^{(\hat{\sigma}_1,\sigma_2)}[(1-\delta_1)\pi(i(\omega;s), j_{s'}) + \delta_1\pi_{1s'+1}|\mathcal{H}_{1s'}].$$

Hence, after the public history $\mathfrak{h}(\omega; s)$, no private history for player 1 (consistent with $F_s^\ddagger$) makes playing action $i(\omega)$ profitable.

References

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.

CRIPPS, M. W., G. J. MAILATH, AND L. SAMUELSON (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72(2), 407–432.

FUDENBERG, D., AND D. K. LEVINE (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.

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.

KALAI, E., AND E. LEHRER (1995): “Subjective Games and Equilibria,” *Games and Economic Behavior*, 8(1), 123–163.