The Dilemma of the Cypress and the Oak Tree

David Rahman*
University of Minnesota

May 27, 2014

Abstract

I study repeated games with mediated communication and frequent actions. I derive Folk Theorems with imperfect public and private monitoring under minimal detectability assumptions. Even in the limit, when noise is driven by Brownian motion and actions are arbitrarily frequent, as long as players are sufficiently patient they can attain virtually efficient equilibrium outcomes, in two ways: secret monitoring and infrequent coordination. Players follow private strategies over discrete blocks of time. A mediator constructs latent Brownian motions to score players on the basis of others' secret monitoring, and gives incentives with these variables at the end of each block to economize on the cost of providing incentives. This brings together the work on repeated games in discrete and continuous time in that, despite actions being continuous, strategic coordination is endogenously discrete. As an application, I show how individual full rank is necessary and sufficient for the Folk Theorem in the Prisoners' Dilemma regardless of whether monitoring is public or private.

JEL Classification: C73, D43, D82.
Keywords: mutual cooperation, mediated communication, frequent actions.

---

*I owe many thanks to David Levine and Joe Ostroy for invaluable comments, the European University Institute for their hospitality during the Fall of 2012 and the National Science Foundation for financial support through Grant No. SES 09-22253.

Contents

"¿Sabes en qué veo que las comiste de tres a tres? En que comía yo dos a dos y callabas."
El Lazarillo de Tormes (Anonymous)

"Let there be spaces in your togetherness, And let the winds of the heavens dance between you. [...]
And stand together, yet not too near together: For the pillars of the temple stand apart,
And the oak tree and the cypress grow not in each other's shadow."
On Marriage (Kahlil Gibran, The Prophet)

1 Introduction

Firms, partners and household members reach complex, dynamic, often informal agreements regarding behavior, coordination and incentives. A basic facet of these relationships involves managing information amongst interested parties. For instance, firms sometimes form trade associations, ostensibly to foment collaboration through regular meetings, as well as standardization. In numerous occasions, these associations have also played the role of information management institutions.[1] This paper studies how such institutions can facilitate mutual cooperation in terms of two canonical channels: (i) they allow players to *secretly monitor* each other, hence only occasionally, which can yield substantial reductions in monitoring costs, and (ii) when incentives require efficiency losses beyond monitoring itself, players can aggregate information better by *coordinating infrequently*.

There is ample empirical evidence for these channels. For instance, on secret monitoring, according to Hexner (1943, p. 95), in a steel manufacturing cartel "[...] more often than one might suppose, infringements of cartel regulations were reported by competing distributions within a few hours [...]." As for infrequent coordination, Marshall and Marx (2012, pp. 126-7) reviewed the 22 major industrial-product cartel decisions of the EC from 2000 through 2005 and concluded that at least half of the cases involved trade associations or other third-party facilitators. The EC Decision on *Amino Acids* (Commission, 2001) reports that a typical cartel member (p. 35) "[...] reported its citric acid sales every month to a trade association, and every year, Swiss accountants audited those figures" for the purpose of sustaining collusion. Moreover (p. 37), a cartel member "[...] further proposed that the producers attend trade association meetings quarterly to adjust their price and sales volumes according to their agreements."

Moreover, these channels may help to identify collusion—apparent cheating behavior among firms is not only observed in practice, but also often goes unpunished, seemingly contradicting standard models of collusion. Bernheim and Madsen (2013, p. 2) argue that "[t]hese unresolved empirical puzzles have important practical implications, in that attorneys for defendant companies often point to evidence of business stealing, and to a purported lack of retaliation, as 'proof' that a cartel is ineffective." The channels above offer a solution to this puzzle: episodes of apparent cheating without punishment are not just tolerable, but possibly essential for collusion to take place in equilibrium.

[1] A recent empirical discussion can be found in Marshall and Marx (2012, Chapter 6), for instance.

To this end, I study repeated games with mediated communication and frequent actions. Players follow *mediated strategies* (Forges, 1986; Myerson, 1986)—a plausible generalization of private strategies that accommodates arbitrary communication.[2] By the revelation principle, these can be represented by a *mediator*[3] who makes confidential, non-binding behavioral recommendations to players. In *communication equilibrium*, everyone follows the mediator's recommendations. Actions are also frequent, where a player's information converges to Brownian motion, for two reasons. First, as Sannikov and Skrzypacz (2010, p. 871) argued, abstracting from the friction of a fixed period length "uncover[s] fundamental principles of [. . .] repeated interactions." It also disciplines significantly how information can be reliably aggregated.[4] Second, to underscore the value of communication, Sannikov and Skrzypacz's (2007; 2010) impossibility results are intuitively overturned.

There is growing interest in games with frequent actions (Sannikov, 2007; Sannikov and Skrzypacz, 2007, 2010; Faingold and Sannikov, 2011; Fudenberg and Levine, 2007, 2009), with a wide range of applications. A key result of this literature is that when information is driven by Brownian motion, "value-burning," or inefficient punishment, is not feasible—this can collapse equilibrium outcomes to the repetition of static equilibria. Perhaps most notably, Sannikov and Skrzypacz (2007) argued that as a result collusion is impossible in a repeated Cournot oligopoly with flexible production.

These papers restrict attention to Nash equilibria in public strategies.[5] Although there is precedent for this solution concept,[6] to classify channels (i) and (ii), in this paper I broaden the notion of equilibrium towards dynamic mechanism design in two steps (see Table 1 below): (i) from public Nash equilibrium to public communication equilibrium, and (ii) from public communication equilibrium to private communication equilibrium. In public Nash equilibrium, strategies only depend on public information. After a mixed strategy, continuation play cannot vary with the mixture's realization: incentives must be independent of actual behavior. In public communication equilibrium, the mediator can keep others' recommendations secret for one period—however brief—and condition

[2]Communication is not new to repeated games (e.g., Compte, 1998; Kandori and Matsushima, 1998; Kandori, 2003; Aoyagi, 2005; Obara, 2009; Tomala, 2009; Harrington and Skrzypacz, 2011), but none of these papers can address frequent actions. Other papers study communication via actions (Ely et al., 2005; Hörner and Olszewski, 2006; Kandori and Obara, 2006; Sugaya, 2010), but also fail with frequent actions.

[3]In principle, this disinterested party can be a machine, a not-necessarily-public randomization device. To be clear, a mediator is a theoretical abstraction that encompasses any communication system. A trade association cannot be interpreted literally as a mediator, but as part of a communication system.

[4]In a simpler context, Rahman (2013) shows that the information aggregation of Abreu et al. (1991) and Compte (1998)—and to some extent Kandori and Matsushima (1998), too—fails with frequent actions.

[5]Recall that a strategy is *public* if it only depends on public histories, otherwise it is *private*.

[6]E.g., Abreu et al. (1990) equate their payoffs to those from pure strategy sequential equilibria.

continuation play on these recommendations as well as any other relevant information. Now incentives can depend on whether a player was secretly monitored. As a result, players can be monitored only seldom, hence at lower monitoring costs. As Rahman (2012b) shows, this restores full collusion in Sannikov and Skrzypacz's (2007) model of oligopoly.

	Nash	*Communication*
Public	Literature	Lemma 1
Private		Lemma 2

Table 1: Strategies and Equilibrium with Frequent Actions

In public communication equilibrium, coordination is so frequent that nobody's private information is useful for more than a period. This may be socially costly, especially in games with frequent actions where a period can be as brief as a nanosecond. Although public communication equilibrium helps to reduce monitoring costs, it can stop short of avoiding the additional costs of value-burning: if, after a deviation, it is impossible to (statistically) identify the culprit, discouraging it requires punishing everyone. With frequent actions and Brownian information, a deviation's noise-to-signal ratio explodes, rendering this punishment infeasible. In private equilibrium, though, even if actions vary frequently, players can temper the noise-to-signal ratio by aggregating private information and synchronizing their histories infrequently, say once a week instead of every nanosecond.

The mediator can manage information as follows. First, he decides on the length of time c during which players' histories may diverge. Each period, he makes secret recommendations, records public signal realizations, and constructs a latent variable Y_{it} for each player i that follows an independent Brownian motion. The drift of Y_i depends on the mediator's recommendations as well as the public signal. Players do not observe the mediator's recommendations to others except at the end of each block, when these are publicly announced together with the latent variables, and another block of length c begins afresh.

I assume a form of conditional identifiability, which implies that these latent variables can be made driftless when players are obedient, but drift downwards when someone disobeys. Everyone is subject to a cutoff that grows linearly in c (so longer blocks have less likely cutoffs). If Y_{ic} is below the cutoff, player i is punished with a loss of continuation value. (I also construct reward schemes with the opposite effect.) A simple cutoff is the lowest possible drift arising from a deviation. This guarantees convexity of punishments, which simplifies the analysis of incentives. However, other cutoffs can motivate cooperation. Punishments turn out to be (approximately) linear in players' discount rate $r > 0$, so they satisfy local self-decomposability, which easily yields a Folk Theorem. Applied to the Prisoners' Dilemma, this result shows how it is always possible to mediate cooperation, except when players cannot learn anything at all about others' behavior.

The main result of the paper is a "Nash-threats" Folk Theorem, which argues that information management can sustain cooperation with frequent actions even with value-burning, as players become unboundedly patient. To this end, I make mostly technical assumptions to easily extend the algorithm of Fudenberg and Levine (1994) to "T-public" communication equilibria. I prove a Folk Theorem conceptually comparable to Compte (1998, Theorem 2) without pairwise identifiability, but with important departures. One key difference is that the scoring rule of Abreu et al. (1991), on which he relies, fails with frequent actions (see Rahman, 2013). I offer an alternative approach below. On the other hand, I do without pairwise identifiability to make value-burning necessary for incentives. To overturn the results of Sannikov and Skrzypacz (2010), I only consider incentives with individual punishments and rewards, hence I only give incentives with value-burning.[7]

The paper considers both public and private monitoring. The most challenging case is public monitoring, though. With private monitoring, the mediator sends recommendations and also asks players to secretly report back their private signals. In this case, it is much easier to keep a score secret from a given player by having it depend on others' reports. With private monitoring, a player does not observe other players' signals, so a player's score is kept secret from that player inasmuch as it depends on others' reported signals.

With public monitoring, however, this secrecy is not available. A key contribution of the paper is to endogenously construct a secret yet informative score with the public signal, which is influenced by actual behavior, and recommendations, which is unaffected by actual behavior within a block. Therefore, the role of recommendations in a player's score is precisely encryption, which is particularly challenging with public monitoring.

This challenge is reflected in my demands on the information structure to establish the Folk Theorems below. With public monitoring, I impose conditional identifiability, which asks that for every public signal and deviation by a player, other players can generate different statistical consequences. This condition is generic if every player has at most as many actions as everyone else put together. With private monitoring (and a standard full support assumption), my demands are much weaker: they are comparable with individual full rank. This is precisely because the challenge of encrypting scoring rules is no longer present with private monitoring—even if monitoring is not conditionally independent.

[7] These incentives are completely different from reputation motives as in Faingold and Sannikov (2011), say. Indeed, take their leading example: a long-run firm with reputation for quality. (With more long-run players, they assume that information has a "product structure," which delivers pairwise identifiability, so value-burning becomes unnecessary.) Myopic short-run players purchase high service levels if they think the long-run player produces high quality. The long-run player does not face uncertainty—he chooses quality over time to manage short-run players' belief-driven service levels. In this paper, though, a long-run player is crucially made uncertain about how he is being evaluated due to moral hazard.

2 Prisoners' Dilemma

Example 1. Two-player repeated Prisoners' Dilemma with imperfect public monitoring.

	C	D
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

Payoffs

	C	D
C	x_2	x_1
D	x_1	x_0

Drifts

The left bi-matrix shows flow payoffs as a function of players' actions. The right matrix shows the drift of a publicly observed Brownian motion X, with law $dX_t = x(a_t)dt + dW_t$, where a_t is the action profile at time t and W is Wiener process. Because volatility is effectively observable, and in line with the literature, assume that actions do not affect the volatility of X. For simplicity, assume that the drift of X only depends on the number of cooperators, not on a cooperator's identity (everything below follows through regardless). Consider discrete-time approximations of this game, where the time between interactions is $\Delta t > 0$, and players have a common discount factor $\delta = e^{-r\Delta t}$ with $r > 0$. In each period, players observe a random walk that converges to the Brownian motion above as $\Delta t \to 0$:

$$X_t = \begin{cases} X_{t-\Delta t} + \sqrt{\Delta t} & \text{with probability } p(a_t) = \frac{1}{2}[1 + x(a_t)\sqrt{\Delta t}], \\ X_{t-\Delta t} - \sqrt{\Delta t} & \text{with probability } q(a_t) = 1 - p(a_t). \end{cases}$$

The choice of a Binomial random walk is not for simplicity: binary per-period signals make incentive provision more challenging in discrete time (Fudenberg et al., 1994) *and* in the continuous-time limit (Fudenberg and Levine, 2009). Below, I estimate the set of communication equilibrium payoffs of this game as $\Delta t \to 0$, and then as $r \to 0$, with the following necessary and sufficient condition for a Folk Theorem, meaning that every non-negative payoff is virtually attainable in equilibrium.

Theorem 1. *The Folk Theorem fails if and only if $x_0 = x_1 = x_2$.*

Theorem 1 gives very weak conditions for a Folk Theorem: as long as $x_0 \neq x_1$ or $x_1 \neq x_2$, it is possible to motivate mutual cooperation as players become patient. Intuitively, this condition says simply that for every unilateral deviation there exists an action profile that makes it statistically detectable, even if the identity of the deviator is not. First of all, necessity is immediate: if $x_0 = x_1 = x_2$ then clearly there is no hope for cooperation, since defecting is completely undetectable, no matter what anybody does. To argue sufficiency, consider two parametric cases corresponding to channels (i) and (ii) above. First, assume that the drift x is monotone in the number of cooperators, to obtain a Folk Theorem in public communication equilibrium that underscores the gains from secret monitoring (Lemma 1). Next, a Folk Theorem in private communication equilibrium is derived when monotonicity fails that highlights the value of infrequent coordination (Lemma 2).

2.1 Secret Monitoring

If x is increasing in the number of cooperators, the secret principal contract of Rahman and Obara (2010) delivers a Folk Theorem with public communication equilibria, as follows.

Lemma 1. *The Folk Theorem holds in public communication equilibrium if $x_2 \geq x_1 > x_0$.*[8]

Proof Sketch. A complete proof appears in Appendix A. Here, I argue that virtually full cooperation is sustainable. Given $\mu \in (0,1)$, with probability $1-\mu$ a mediator secretly recommends cooperation (and defection with probability μ) independently to each player. Recommendations are publicly announced at the end of each period, after X is realized. Current-period expected payoff is $1-\mu$. By Theorem 4.1 of Fudenberg et al. (1994), it suffices to show that this μ is enforceable with respect to tangent hyperplanes, i.e., there exist budget-balanced transfers that make μ incentive compatible. Let continuation values change depending on whether X jumps up (labeled $+1$) or down (labeled -1) as well as the mediator's recommendations as follows:

$+1$	C	D
C		$+w, -w$
D	$-w, +w$	

-1	C	D
C		$-w, +w$
D	$+w, -w$	

Change in continuation values after $+1$ and -1

If the mediator asks a player to cooperate, incentive compatibility requires that

$$(1-\delta)(1-2\mu) + \delta\mu w(p_1 - q_1) \geq (1-\delta)2(1-\mu) + \delta\mu w(p_0 - q_0).$$

Rearranging and substituting for p and q yields, equivalently, $\delta\mu w(x_1 - x_0)\sqrt{\Delta t} \geq 1-\delta$. Since $\delta = e^{-r\Delta t}$ and $1-\delta \leq r\Delta t$, this follows from the following inequality:

$$\delta\mu w \geq \frac{r\sqrt{\Delta t}}{x_1 - x_0}. \tag{1}$$

Similarly, for a player asked to defect, incentive compatibility requires that

$$(1-\delta)2(1-\mu) + \delta w(1-\mu)(q_1 - p_1) \geq (1-\delta)(1-2\mu) + \delta w(1-\mu)(q_2 - p_2),$$

which, rearranging as in the previous derivation, follows from $w \geq 0$ and hence is implied by (1). Intuitively, cooperating when asked to defect lowers utility this period and the probability of a down jump next period, since $x_2 > x_1$. Choose w such that (1) holds with equality, so w is increasing in both r and Δt, (almost) linear in r, and the correlated strategy

[8] Contrast this with Sannikov and Skrzypacz (2010, Example 1), where $(x_0, x_1, x_2) = (1, 5, 8)$. There, they show that welfare cannot exceed 1 in public Nash equilibrium when $\Delta t \to 0$.

above is enforceable in the welfare direction $(1,1)$. Given a smooth set in the interior of the feasible, individually rational payoffs, its boundary point with tangent vector $(1,1)$ exhibits local self-decomposability (LSD) for some $(r,\Delta t)$ by Theorem 4.1 of Fudenberg et al. (1994). If $\underline{w}$ yields LSD at $(r,\underline{\Delta})$ as above then so does $w = \underline{w}\sqrt{\Delta t/\underline{\Delta}}$ at $(r,\Delta t)$ for any $\Delta t < \underline{\Delta}$. Finally, $w \to 0$ at rate r as $r \to 0$—the Folk Theorem only requires convergence faster than $\sqrt{r}$. See Figure 1 below for geometric intuition. $\square$

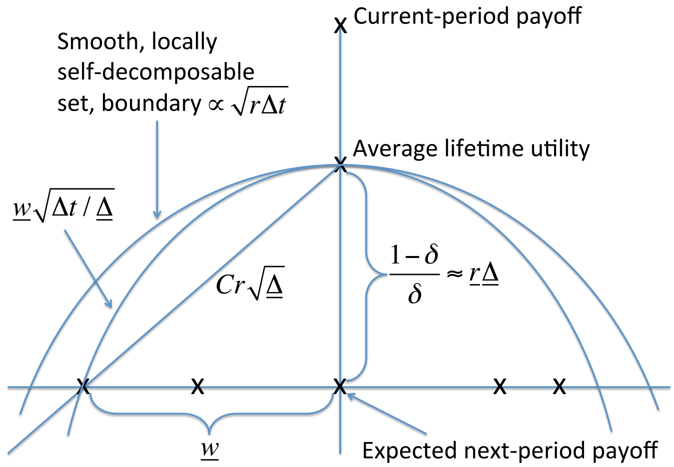

Figure 1: Frequent-actions versus discrete-time Folk Theorem

In the proof of Lemma 1, $x_2 \geq x_1 > x_0$ is used to establish enforceability with respect to every tangent hyperplane. Reversing the roles of up and down jumps, the same proof follows verbatim if $x_0 > x_1 \geq x_2$. This monotonicity corresponds to first-order stochastic dominance in the literature on oligopoly with noisy prices, from Green and Porter (1984) and Abreu et al. (1986) to Sannikov and Skrzypacz (2007) and Harrington and Skrzypacz (2011). Monotonicity permits identifying obedient agents (Rahman and Obara, 2010), and assigning continuation values that avoid "value-burning" (Sannikov and Skrzypacz, 2010). To see why, let $x_2 \geq x_1 \not> x_0$ (but not $x_0 = x_1 = x_2$) and consider the following profile of deviations: a player asked to cooperate defects with probability α, whereas if asked to defect he cooperates with probability β such that $0 \leq \alpha, \beta \leq 1$ and $\beta/\alpha = (p_0 - p_1)/(p_2 - p_1)$.[9] For any correlated strategy, it is impossible to identify the deviator after a unilateral deviation.[10] As a result, Lemma 1 generally fails: since it could have been anyone, every player must be punished on-path. This leads to inefficiency, even as players become patient. In other words, value-burning is unavoidable without monotonicity.

[9]For instance, if $x_2 = x_0 = 1$ and $x_1 = 0$, let $\alpha = 1$ and $\beta = 1$: disobedience with probability one.

[10]Since deviations are identical, a deviator cannot be identified from (C,C) or (D,D). Given (C,D), if player 1 defects then the distribution of signals changes by $\alpha[(p_0,q_0)-(p_1,q_1)]$, just as for player 2, since $\beta[(p_1,q_1)-(p_2,q_2)] = \alpha[(p_0,q_0)-(p_1,q_1)]$. By symmetry, the claim now follows.

2.2 Infrequent Coordination

Such inefficiency will now be overcome by manipulating the arrival of recommendations, yielding a Folk Theorem in private communication equilibrium assuming that x satisfies

$$x_2 - x_1 \neq x_1 - x_0, \qquad \text{or, equivalently,} \qquad x_0 + x_2 \neq 2x_1. \tag{2}$$

The only vectors x excluded by (2) exhibit $x_2 - x_1 = x_1 - x_0$. If $x_2 - x_1 = x_1 - x_0 > 0$ then $x_2 > x_1 > x_0$, and by Lemma 1 a Folk Theorem obtains. Similarly, a Folk Theorem holds if $x_2 - x_1 = x_1 - x_0 < 0$ by reversing the roles of up and down jumps. Hence, the Folk Theorem fails only if $x_2 - x_1 = x_1 - x_0 = 0$, proving Theorem 1.

Let us construct T-period blocks such that $T = \lfloor c/\Delta t \rfloor$, where c is a constant representing the length of *calendar time* of each block. To sustain a nearly efficient payoff, players face a punishment scheme that depends on both the mediator's secret recommendations over the T-period block and the public signals. At the end of the T-period block all previous recommendations are made public to solve for T-public communication equilibria. This delays the arrival of inefficient punishments and increases the number of signals contingent on which to trigger said punishments, which tempers the noise to signal ratio.

Lemma 2. *The Folk Theorem also holds if $x_2 - x_1 \neq x_1 - x_0$.*

Proof Sketch. This result follows from Theorem 2 below, proved in Appendix A. Here I argue that virtually full cooperation is sustainable. Without loss, let $x_0 + x_2 > 2x_1$. The case $x_0 + x_2 < 2x_1$ is handled similarly by reversing the roles of up and down jumps. The mediator generates a latent variable Y_i for each player i that follows a random walk representation of Brownian motion, with drift determined by the mediator's recommendations and the public signals. Players do not observe these latent variables except every c units of time. The mediator secretly recommends cooperation to each player with probability $1 - \mu$ and defection with probability μ, for some small $\mu > 0$. Players observe only their own recommendations. Recommendations are IID across players and time throughout each block. At time t, the mediator makes a profile of recommendations a_t, each player takes a possibly different action, and the publicly observed variable X_t realizes. Let $\omega_t = +1$ if X_t jumps up and -1 if it jumps down. Take the first block of length c. The latent variable Y_i starts at $Y_{i0} = 0$. After (a_t, ω_t) realizes, the mediator performs a secret Bernoulli trial with failure probability $\xi_i(a_t, \omega_t)$, for some scoring rule ξ that drives the law of motion of Y_i:

$$Y_{it} = \begin{cases} Y_{it-\Delta t} + \sqrt{\Delta t} & \text{with probability } \zeta_i(a_t, \omega_t) = 1 - \xi_i(a_t, \omega_t), \\ Y_{it-\Delta t} - \sqrt{\Delta t} & \text{with probability } \xi_i(a_t, \omega_t). \end{cases}$$

Assuming (2), there exists a scoring rule ξ such that deviations increase failure rates, even after obedience failure is still possible, and ξ enforces the correlated strategy defined by μ.

The following scoring rule works: $\xi_i(a,\omega) = \frac{1}{2}$ for all (i,a,ω) *except*

$$\xi_i(C_i, C_{-i}, +1) = \tfrac{1}{2}\left[1 - \frac{\mu}{1-\mu}\frac{p_1}{p_2}\right] \quad \text{and} \quad \xi_i(C_i, D_{-i}, +1) = 1,$$

where $\mu > 0$ and $\Delta t > 0$ are small enough that $\xi_i(a,\omega) \in [0,1]$ is a probability.

Observe that $\xi_i(C_i, D_{-i}, +1) > \frac{1}{2} > \xi_i(C_i, C_{-i}, +1)$: an increase in X_t increases the failure rate if i's opponent was recommended to defect and lowers it otherwise. Notice that, on the path of play, player i's failure rate conditional on his information equals $\frac{1}{2}$, so the latent variable Y_i follows a random walk without drift. Indeed, after i was recommended D_i or X jumped down this is clear by construction of ξ. Moreover, the probability that Y_i jumps down given that X jumped up, i was recommended C_i and he obeyed equals

$$\frac{(1-\mu)\frac{1}{2}\left[1 - \frac{\mu}{1-\mu}\frac{p_1}{p_2}\right]p_2 + \mu p_1}{(1-\mu)p_2 + \mu p_1} = \frac{\frac{1}{2}[(1-\mu)p_2 + \mu p_1]}{(1-\mu)p_2 + \mu p_1} = \tfrac{1}{2}.$$

If player i defected when asked to cooperate, this conditional probability increases to

$$\pi^{**} = \frac{(1-\mu)\frac{1}{2}\left[1 - \frac{\mu}{1-\mu}\frac{p_1}{p_2}\right]p_1 + \mu p_0}{(1-\mu)p_1 + \mu p_0} = \tfrac{1}{2}\left[1 + \mu\frac{(p_0p_2 - p_1^2)/p_2}{(1-\mu)p_1 + \mu p_0}\right].$$

As $\Delta t \to 0$, since $p_2 \approx (1-\mu)p_1 + \mu p_0 \approx \frac{1}{2}$, this conditional probability is close to $\frac{1}{2}[1 + \mu(x_0 + x_2 - 2x_1)\sqrt{\Delta t}]$. Hence, the largest possible conditional failure *drift* is

$$z^{**} = \mu(x_0 + x_2 - 2x_1).$$

Unconditionally, that is, before the realization of X, the probability of failure equals

$$\pi^* = (1-\mu)\tfrac{1}{2}\left(\left[1 - \frac{\mu}{1-\mu}\frac{p_1}{p_2}\right]p_1 + q_1\right) + \mu(p_0 + \tfrac{1}{2}q_0) = \tfrac{1}{2}\left[1 + \mu(p_0p_2 - p_1^2)/p_2\right].$$

Hence, the unconditional, or prior, failure drift is similarly derived from π^* to be

$$z^* = \tfrac{1}{2}\mu(x_0 + x_2 - 2x_1).$$

At the end of the block, that is, at time c, the mediator publicly reveals the entire path of recommendations as well as each Y_i, and the next block begins afresh. If the value of Y_i is below some threshold $Y^{**} < 0$ at time c, then player i pays a penalty w of continuation value. For fixed $\Delta t > 0$, if Y_i exhibited $t^{**} = \lfloor(1-\pi^{**})(T-1)\rfloor$ successes or fewer along the block then punishment ensues. The threshold Y^{**} is related to t^{**} through the random walk representation.[11] As $\Delta t \to 0$, the threshold is linear in the block length c:

$$Y^{**} \;\to\; -z^{**}c \quad \text{as} \quad \Delta t \to 0.$$

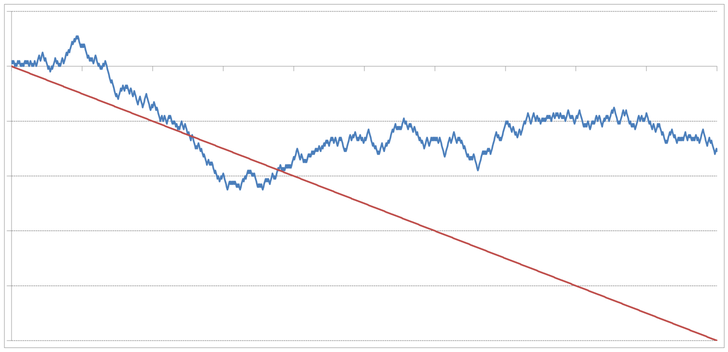

Figure 2: Linear threshold (red) versus latent variable (blue) over time

In equilibrium, Y_i is a driftless Brownian motion, so $Y_{ic} - Y_{i0} = Y_{ic} \sim N(0, c)$ and i is punished with probability $1-\Phi(z^{**}\sqrt{c})$, where Φ is the standard normal CDF. Player i only affects the drift of Y_i by deviating, hence the mean of Y_{ic}. If player i is asked to cooperate in the first period but defects instead, he is punished with probability $1-\Phi((z^{**}c-z^*\Delta t)/\sqrt{c})$.

To discourage this single deviation, punishment costs must outweigh deviation gains:

$$1 - e^{-r\Delta t} \leq e^{-rc} w[\Phi(z^{**}\sqrt{c}) - \Phi((z^{**}c - z^*\Delta t)/\sqrt{c})]. \tag{3}$$

Since $1-e^{-x} \leq x$ for $x \geq 0$, the left-hand side is estimated by $r\Delta t$, and since $\Phi(z)$ is concave for $z \geq 0$, the right-hand side is estimated by $e^{-rc}w\varphi(z^{**}\sqrt{c})z^*\Delta t/\sqrt{c}$. Substituting these estimates, multiplying both sides by $c/\Delta t$ and rearranging, (3) follows from

$$w \geq \frac{rce^{rc}}{\varphi(z^{**}\sqrt{c})z^*\sqrt{c}}.$$

If w makes this inequality bind then every deviation is discouraged. To see this, pick any partial history of behavior and observations, h_i^t. Relative to obedience, the mean of Y_{ic} may have decreased by some amount, say θ. By construction of ξ, however, θ cannot exceed $z^{**}c$, the latter being c times the largest possible conditional failure drift. With the same logic as for (3), a single deviation after h_i^t is discouraged if $w \geq rce^{rc}/[\varphi((z^{**}c - \theta)/\sqrt{c})z^*\sqrt{c}]$. This clearly holds since $\varphi((z^{**}c - \theta)/\sqrt{c}) \leq \varphi(z^{**}\sqrt{c})$ for all $\theta \in [0, z^{**}c]$, so at any partial history, every one-step deviation is unprofitable. Hence, *every* deviation is discouraged.

Since w is approximately linear in r for every c, it diminishes faster than $\sqrt{r}$ as $r \to 0$, so punishments are self-decomposable if players are patient. Each player's lifetime expected utility solves $v = (1 - e^{-rc})(1 - \mu) + e^{-rc}[\Phi(z^{**}\sqrt{c})v + (1 - \Phi(z^{**}\sqrt{c}))(v - w)]$, so

$$v = 1 - \mu - \frac{rc}{1 - e^{-rc}} \frac{1 - \Phi(z^{**}\sqrt{c})}{\varphi(z^{**}\sqrt{c})z^*\sqrt{c}}.$$

Finally, as $r \to 0$, this holds for every c, and $rc/(1 - e^{-rc}) \to 1$. As $c \to \infty$, the normal hazard rate explodes linearly, so $v \to 1 - \mu$, and, since $\mu > 0$ was arbitrary, $v \to 1$. $\square$

[11]Roughly, Y_{ic} equals Y^{**} after t^{**} jumps up and $T - t^{**}$ jumps down, each jump having length $\sqrt{\Delta t}$. Thus, $Y^{**} \approx t^{**}\sqrt{\Delta t} - (T - t^{**})\sqrt{\Delta t} \approx (1 - 2\pi^{**})T\sqrt{\Delta t}$.

Mutual cooperation requires so little in Theorem 1 because players just have two actions. With more actions, more is required for a Folk Theorem, but the strategic flavor is similar. On the other hand, with private monitoring (and full support) the requirements for a Folk Theorem are much weaker, reducing to individual identifiability. See Section 6.

The threshold Y^{**} above is clearly not the only one that can induce mutual cooperation. I chose it to yield a particularly transparent derivation of incentive compatibility. Finding optimal thresholds and strategies is an interesting open question for further research.

The mediator may be dispensable. Instead of taking recommendations, players may simply report their *intended action* before playing it. This is reminiscent of but not identical to Kandori's (2003) contracts. To reveal his intentions, a player must be indifferent over all reports. Nevertheless, Section 6 shows that reporting intentions can sometimes—yet decidedly not always—be a good substitute for a mediator. Finally, even if a mediator cannot be dispensed with, it may be possible to decentralize it with plain conversation, as argued by Forges (1986, 1990). She required four or more players, though.

3 Assumptions

3.1 Payoffs

Consider a repeated game with imperfect monitoring in discrete time. Each stage of the game is indexed by $\tau \in \{1, 2, \ldots\}$, with $\Delta t > 0$ the length of time between stages. The calendar date of each stage is $t \in \{\Delta t, 2\Delta t, \ldots\}$. The stage game is repeated every period; it consists of a finite set $I = \{1, \ldots, n\}$ of players, a finite set A_i of actions for each player $i \in I$, where $A = \prod_i A_i$, and a function $u : I \times A \to \mathbb{R}$, where $u_i(a)$ denotes the utility flow to player i from action profile a. Players have a common discount factor $\delta = e^{-r\Delta t}$ with $r\Delta t > 0$. The utility to player i from a sequence of action profiles $a^\infty = (a_1, a_2, \ldots)$ equals

$$(1 - \delta) \sum_{\tau=1}^{\infty} \delta^{\tau-1} u_i(a_\tau).$$

Let $U = \{u(\mu) = \sum_a u(a)\mu(a) : \mu \in \Delta(A)\}$ be the convex hull of stage-game payoff vectors. Given a correlated strategy $\mu \in \Delta(A)$, let

$$\underline{u}_i(\mu) = \max_{\sigma_i : A_i \to \Delta(A_i)} \sum_{(a,b_i)} u_i(b_i, a_{-i})\mu(a)\sigma_i(b_i|a_i),$$

and write $\underline{u}(\mu) = (\underline{u}_1(\mu), \ldots, \underline{u}_n(\mu))$. Player i's correlated *minmax* value is given by

$$\underline{u}_i = \min_{\mu \in \Delta(A)} \underline{u}_i(\mu).$$

Write $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ for the vector of such values across all players. The set of *feasible, individually rational payoffs* is denoted by $\underline{U} = \{u \in U : u \geq \underline{u}\}$.

Assumption 1 (Payoffs)**.** (a) The set $\underline{U}$ has dimension n. (b) The stage game has a strictly inefficient correlated equilibrium, with payoff profile u_0. Let $U_0 = \{u \in \underline{U} : u \geq u_0\}$.

This is standard. Assumption 1(a) says that in principle players can be punished individually and 1(b) that there is room for everyone's improvement beyond static equilibrium.

3.2 Probabilities

Let Ω be a finite set of signals observed every period, and assume that $|\Omega| > 1$. In case of public monitoring, $\omega \in \Omega$ is observed in every period by every player. In case of private monitoring, $\Omega = \prod_{i=1}^{n} \Omega_i$ and each player i observes some $\omega_i \in \Omega_i$ per period. Let $\Pr(\omega|a)$ be the probability that $\omega \in \Omega$ realizes at the end of each period when $a \in A$ was played in that same period. In general, the matrix Pr also depends on Δt, as is discussed later.

Assumption 2 (Full support)**.** There exists $\underline{\pi} > 0$ such that $\Pr(\omega|a) \geq \underline{\pi}$ for all $(a, \omega, \Delta t)$.

This assumptions is standard. It is useful to distinguish public and private monitoring. I now define my main notions of identifiability. The first applies to both public and private monitoring, and is stronger than the second in case of private monitoring.

Definition 1. The matrix Pr exhibits *conditional identifiability* (CI) if

$$\Pr(\omega|a_i, \cdot) \notin \text{cone}\{\Pr(\omega|b_i, \cdot) : b_i \neq a_i\} \quad \forall (i, a_i, \omega).$$

In case of private monitoring, Pr exhibits *unconditional identifiability* (UI) if

$$\Pr(\omega_i, \cdot|a_i, \cdot) \notin \text{cone}\{\Pr(\omega_i', \cdot|a_i', \cdot) : (a_i', \omega_i') \neq (a_i, \omega_i)\} \quad \forall (i, a_i, \omega_i).$$

Conditional identifiability is stronger than a conic version of individual full rank, which requires only that $\Pr(\cdot|a_i, \cdot) \notin \text{cone}\{\Pr(\cdot|b_i, \cdot) : b_i \neq a_i\}$ for all (i, a_i).[12] Nevertheless, it is equally silent about the identity of a deviator. It is consistent with some strongly symmetric conditional distributions, unlike pairwise identifiability, for instance. On the other hand, in case of private monitoring, unconditional identifiability is not stronger than the conic version of individual full rank. In fact, the two notions coincide.

CI is generic if every player has at most as many actions as everyone else put together: $|A_i| \leq |A_{-i}|$ for all i, regardless of the (finite) number of signals. Moreover, UI is generic if $|\Omega_i| > 1$ and $|A_i \times \Omega_i| \leq |A_{-i} \times \Omega_{-i}|$ for all i. (See Proposition 4 for full characterization.) Intuitively, this condition takes into account the fact that, with private monitoring, it must be possible to detect not only when a player disobeys, but also when he misreports his observations. Neither of these two genericity conditions seems unnatural or stringent.

[12]This is still significantly stronger than convex independence, rather than the conic version above. Convex independence is necessary but not sufficient for the Folk Theorem below.

Example 2. Recall the Prisoners' Dilemma with public monitoring of Section 2. Let $q_2 = \Pr(-1|C,C)$, $q_1 = \Pr(-1|C,D) = \Pr(-1|D,C)$ and $q_0 = \Pr(-1|D,D)$. Assume that $0 < q_2 < q_1 < 1$. Conditional identifiability holds whenever $q_1 \geq q_0$. Indeed, by definition it fails if and only if either $q_2/q_1 = q_1/q_0$ or $(1-q_2)/(1-q_1) = (1-q_1)/(1-q_0)$ or both, which clearly requires that $q_1 < q_0$, since $q_2 < q_1$.

For interpretation, pick any player i, profile $a \in A$ of mediator recommendations and mixed strategy $\sigma_i \in \Delta(A_i)$ with $\sigma_i(a_i) < 1$. The ratio $\sigma_i(a_i)\Pr(\omega|a)/\Pr(\omega|\sigma_i, a_{-i})$, where $\Pr(\omega|\sigma_i, a_{-i}) = \sum_{b_i} \sigma_i(b_i)\Pr(\omega|b_i, a_{-i})$, equals the posterior probability that player i obeys the mediator given the mediator's information: his recommendations and the public signal. Conditional identifiability holds if and only if for every signal ω there exist two profiles, a_{-i} and b_{-i}, that give the mediator different posterior beliefs of whether or not i played a_i, that is, $\frac{\sigma_i(a_i)\Pr(\omega|a)}{\Pr(\omega|\sigma_i,a_{-i})} \neq \frac{\sigma_i(a_i)\Pr(\omega|a_i,b_{-i})}{\Pr(\omega|\sigma_i,b_{-i})}$. Otherwise, player i could infer the mediator's beliefs. Therefore, if player i does not observe whether a_{-i} or b_{-i} was recommended, the mediator may sustain i's belief that punishment is possible even when the mediator knows it is not.

With private monitoring, the same interpretation applies to unconditional identifiability when a player i's deviation $\hat{\sigma}_i \in \Delta(A_i \times \Omega_i)$ now includes misreporting of privately observed signals. Since we will be studying communication equilibria, in case of private monitoring every player i will be assumed to report the realization ω_i of her signal to a mediator. In this case, let $\Pr(\omega_{-i}|\hat{\sigma}_i, a_{-i}) = \sum_{(a_i',\omega_i')} \hat{\sigma}_i(a_i', \omega_i')\Pr(\omega_i', \omega_{-i}|a_{-i}, a_i')$.

To simplify notation, with public monitoring let $\mathrm{Lr}(\omega|a,\sigma_i) = \Pr(\omega|a)/\Pr(\omega|\sigma_i, a_{-i})$ and

$$\Delta\mathrm{Lr}(\omega|a_i,\sigma_i) = \max_{(a_{-i},b_{-i})} \mathrm{Lr}(\omega|a,\sigma_i) - \mathrm{Lr}(\omega|a_i, b_{-i}, \sigma_i).$$

With private monitoring, let $\mathrm{Lr}(\omega|a,\hat{\sigma}_i) = \Pr(\omega|a)/\Pr(\omega_{-i}|\hat{\sigma}_i, a_{-i})$ and write

$$\Delta\mathrm{Lr}(\omega_i|a_i,\hat{\sigma}_i) = \max_{(a_{-i},\omega_{-i},a_{-i}',\omega_{-i}')} \mathrm{Lr}(\omega|a,\hat{\sigma}_i) - \mathrm{Lr}(\omega_i, \omega_{-i}'|a_i, a_{-i}', \hat{\sigma}_i).$$

Lemma 3. *CI fails if and only if there exists $(i, a_i, \omega, \sigma_i)$ such that $\sigma_i(a_i) < 1$ and $\Delta\mathrm{Lr}(\omega|a_i,\sigma_i) = 0$. With private monitoring, UI fails if and only if there exists $(i, a_i, \omega_i, \hat{\sigma}_i)$ such that $\hat{\sigma}_i(a_i, \omega_i) < 1$ and $\Delta\mathrm{Lr}(\omega_i|a_i,\hat{\sigma}_i) = 0$.*

3.3 Drifts

For a Folk Theorem with frequent actions, I require a "closed" version of identifiability as $\Delta t \to 0$. In other words, I assume that deviations are detectable by at least order $\sqrt{\Delta t}$.

Definition 2. Pr exhibits *closed conditional identifiability* ($\overline{\mathrm{CI}}$) if $(\varepsilon, \varepsilon')$ exist such that

$$\inf_{\Delta t > 0} \frac{\Delta\mathrm{Lr}(\omega|a_i,\sigma_i)}{\sqrt{\Delta t}} > \varepsilon > 0 \qquad \forall (i, a_i, \omega, \sigma_i) \text{ s.t. } \sigma_i(a_i) < 1 - \varepsilon' < 1.$$

With private monitoring, *closed unconditional identifiability* ($\overline{\mathrm{UI}}$) holds if, for some $(\varepsilon, \varepsilon')$,

$$\inf_{\Delta t > 0} \frac{\Delta \mathrm{Lr}(\omega_i | a_i, \hat{\sigma}_i)}{\sqrt{\Delta t}} > \varepsilon > 0 \qquad \forall (i, a_i, \omega_i, \hat{\sigma}_i) \text{ s.t. } \hat{\sigma}_i(a_i) < 1 - \varepsilon' < 1.$$

These conditions restrict the set of stochastic processes for which I prove a Folk Theorem. They are rather general, satisfied by a lot of processes. For instance, many random walks that converge to Brownian motion satisfy these conditions. Other stochastic processes satisfy these conditions, too, but arguably the prototypical example to be studied first is Brownian motion. Closed identifiability simply asks that the extent to which the mediator can separate an action from its deviation does not vanish as fast or faster than $\sqrt{\Delta t}$.

Example 3. To illustrate, the main example below is a natural extension of the discrete-time model with imperfect public monitoring in Example 1—a Binomial random walk that converges to Brownian motion as actions become arbitrarily frequent. Let $\Omega = \{\pm 1\}$. Given an action profile $a \in A$, a function $x : A \to \mathbb{R}$ and $\Delta t > 0$ sufficiently small, let

$$\begin{aligned} \Pr(+1|a) &= \tfrac{1}{2}[1 + x(a)\sqrt{\Delta t}] &&=: \ p(a) \quad \text{and} \\ \Pr(-1|a) &= \tfrac{1}{2}[1 - x(a)\sqrt{\Delta t}] &&=: \ q(a) \ = \ 1 - p(a). \end{aligned}$$

Consider the following Binomial random walk X starting at $X_0 = 0$. For all $t > 0$, let $X_t = X_{t-\Delta t} + \varepsilon_t$, where $\{\varepsilon_t\}$ is a sequence of independent Bernoulli trials with success probability $p(a_t)$. Success means that $\varepsilon_t = \sqrt{\Delta t}$, so X jumps up, whereas failure implies $\varepsilon_t = -\sqrt{\Delta t}$, so X jumps down. As $\Delta t \to 0$, X converges to a Brownian motion with law

$$dX_t = x(a_t)dt + dW_t,$$

where W corresponds to Wiener process. Thus, x describes the drift of X.

Since the probability matrix above depends on Δt, it is convenient to find a condition on the drift function x that guarantees conditional identifiability for small Δt.

Definition 3. The drift function x exhibits *conditional identifiability* (CI-x) if

$$x(a_i, \cdot) \notin \mathrm{conv}\{x(b_i, \cdot) : b_i \neq a_i\} + L_{\mathbf{1}} \quad \forall (i, a_i), \tag{4}$$

where "conv" stands for convex hull and $L_{\mathbf{1}} = \{\alpha(1, \ldots, 1) \in \mathbb{R}^{A_{-i}} : \alpha \in \mathbb{R}\}$.

I will write conditional identifiability "in probabilities" versus "in drifts" when necessary.

Lemma 4. *If x exhibits conditional identifiability then for some $\underline{\Delta} > 0$ and all $\Delta t \in (0, \underline{\Delta})$, the probability matrix that parametrizes X at Δt exhibits conditional identifiability.*

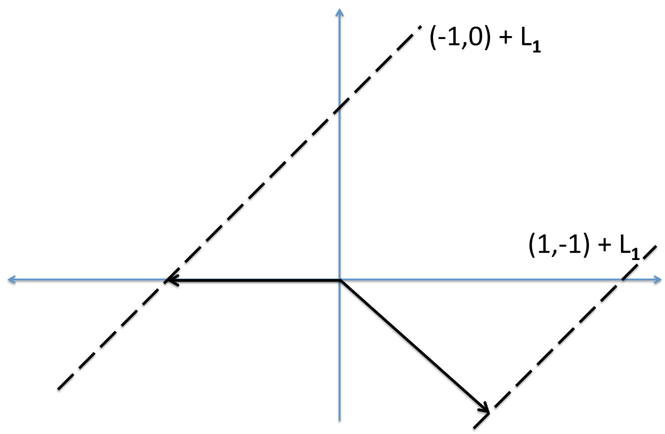

Figure 3: Conditional identifiability in Example 1 when $(x_0, x_1, x_2) = (0, -1, 1)$

Lemma 4 shows that conditional identifiability of the limiting stochastic process implies conditional identifiability in probabilities of its random walk representation sufficiently close to it. For intuition, recall the Prisoners' Dilemma of Example 1. Figure 3 below shows that (4) holds when $(x_0, x_1, x_2) = (0, -1, 1)$. However, (4) is sufficient but not necessary for conditional identifiability in probabilities given all small Δt, as the next example shows.

Example 4. With reference to Example 1, let $x_1 = \frac{1}{2}(x_0 + x_2)$ but $x_0 \neq x_2$. Clearly, (4) fails, yet conditional identifiability holds for all Δt. Indeed, conditional identifiability requires that both $p_2/p_1 \neq p_1/p_0$ and $q_2/q_1 \neq q_1/q_0$. This is easily seen to be equivalent to

$$|2x_1 - (x_0 + x_2)| \neq \left|x_1^2 - x_0 x_2\right| \sqrt{\Delta t}.$$

By hypothesis, the left-hand side of this inequality equals zero and the right-hand side does not, so conditional identifiability holds for all $\Delta t > 0$ even though (4) fails. On the other hand, the right-hand side tends to zero as $\Delta t \to 0$, so conditional identifiability subsides.

Proposition 1. *CI-x fails if and only if $\overline{CI}$ fails, that is, there exists some $(i, a_i, \sigma_i, \omega)$ such that $\sigma_i(a_i) < 1$ and*

$$\frac{\Delta \mathrm{Lr}(\omega | a_i, \sigma_i)}{\sqrt{\Delta t}} \to 0 \quad \text{as} \quad \Delta t \to 0.$$

Intuitively, Proposition 1 says that CI-x fails whenever the mediator's informational advantage—in terms of having several possible posterior beliefs of whether player i obeyed the mediator given the mediator's information—deteriorates faster than $\sqrt{\Delta t}$. This is important because, although by Lemma 4 conditional identifiability in drifts is not necessary for conditional identifiability in probabilities for small Δt, the Folk Theorem below relies on the full strength of CI-x (or more generally $\overline{\mathrm{CI}}$) beyond CI for arbitrarily small Δt.

Assumption 3 (Drifts)**.** $\overline{\mathrm{CI}}$ ($\overline{\mathrm{UI}}$) holds under public (private) monitoring.

4 Communication Equilibrium

In communication equilibrium (Myerson, 1986; Forges, 1986), a disinterested mediator sends non-binding messages to players, who use them to make inferences about others' behavior and best-respond. With private monitoring, players also send messages to the mediator after making their observations. Below, I formally define public and private communication equilibrium. For simplicity, the first two sections assume public monitoring. The last section shows how the construction of the first two sections applies to private monitoring with arguably minor adjustments. I suggest skipping this section to those familiar with public and communication equilibrium.

To classify equilibria, I decompose the mediator's messages into private and public ones. Let A_i be the set of private *recommendations* that the mediator can send to i and $\mathcal{A}$ the finite set of possible public *announcements*. At the beginning of any period τ, the mediator makes a private, non-binding recommendation to every i to play $a_{i\tau} \in A_i$. A public signal $\omega_\tau \in \Omega$ realizes depending on what players actually played. Finally, the mediator sends a public announcement $\alpha_\tau \in \mathcal{A}$ to everyone. Let $H_0^\tau = (A \times \Omega \times \mathcal{A})^{\tau-1}$ be the set of partial histories for the mediator (thus $H_0^1 = \{\emptyset\}$), consisting of recommendations, public signals and announcements. The set of all such partial histories is $H_0 = \bigcup_\tau H_0^\tau$.

Definition 4. A *communication mechanism*, or *mediated strategy*, is a pair $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2)$, where $\tilde{\mu}_1(a_t|a^{\tau-1}, \omega^{\tau-1}, \alpha^{\tau-1})$ is the conditional probability that the mediator privately recommends $a_{i\tau}$ to every player i given $(a^{\tau-1}, \omega^{\tau-1}, \alpha^{\tau-1})$, and $\tilde{\mu}_2(\alpha_\tau|a^\tau, \omega^\tau, \alpha^{\tau-1})$ is the conditional probability that the subsequent public announcement is α_τ.

Notation. I will write $\tilde{\mu}$ to describe mediated strategies of the repeated game, and write $\mu \in \Delta(A)$ for a correlated strategy of the stage game.

A mediated strategy $\tilde{\mu}$ induces a probability distribution on each H_0^τ as follows:

$$\Pr(a^\tau, \omega^\tau, \alpha^\tau|\tilde{\mu}) = \prod_{s=1}^{\tau} \tilde{\mu}_1(a_s|a^{s-1}, \omega^{s-1}, \alpha^{s-1})\tilde{\mu}_2(\alpha_s|a^s, \omega^s, \alpha^{s-1}) \Pr(\omega_s|a_s).$$

The utility to player i from the mediated strategy $\tilde{\mu}$ is therefore given by

$$U_i(\tilde{\mu}) = (1-\delta) \sum_{(\tau, a^\tau, \omega^\tau, \alpha^\tau)} \delta^{\tau-1} u_i(a_\tau) \Pr(a^\tau, \omega^\tau, \alpha^\tau|\tilde{\mu}).$$

Let $H_i^\tau = (A_i \times A_i \times \Omega \times \mathcal{A})^{\tau-1}$ be the set of partial histories for player i, with typical element $h_i^\tau = (a_i^{\tau-1}, b_i^{\tau-1}, \omega^{\tau-1}, \alpha^{\tau-1})$, where $a_i^{\tau-1}$ is the vector of recommendations to player i from periods 1 to $\tau - 1$, $b_i^{\tau-1}$ is the vector of actions taken by player i, $\omega^{\tau-1}$ is the vector of signal realizations and $\alpha^{\tau-1}$ is the vector of public announcements by the mediator. Let $H_i = \bigcup_t H_i^\tau$ be the set of all partial histories for player i.

For any player i, a unilateral *deviation* from $\tilde{\mu}$ is a function $\sigma_i : H_i \to M(A_i)$, where $M(A_i) = \{A_i \to \Delta(A_i)\}$ is the set of recommendation-contingent mixed strategies and $\sigma_i(b_{i\tau}|a_{i\tau}, h_i^\tau)$ is interpreted as the probability that player i plays $b_{i\tau}$ if $a_{i\tau}$ is recommended in period τ when his private history is h_i^τ. The induced probability that $(a^\tau, \omega^\tau, \alpha^\tau, b_i^\tau)$ occurs if everyone else is honest and obedient except for i, who deviates accord to σ_i, is

$$\Pr(a^\tau, \omega^\tau, \alpha^\tau, b_i^\tau|\tilde{\mu}, \sigma_i) = \Pr(a^\tau, \omega^\tau, \alpha^\tau|\tilde{\mu}) \prod_{s=1}^{\tau} \sigma_i(b_{is}|a_i^s, b_i^{s-1}, \omega^{s-1}, \alpha^{s-1}) \frac{\Pr(\omega_s|b_{is}, a_{-is})}{\Pr(\omega_s|a_s)}.$$

Therefore, the utility to player i from a unilateral deviation σ_i can be written as

$$U_i(\tilde{\mu}|\sigma_i) = (1-\delta) \sum_{(\tau, a^\tau, b_i^\tau, \omega^\tau, \alpha^\tau)} \delta^{\tau-1} u_i(b_{i\tau}, a_{-i\tau}) \Pr(a^\tau, \omega^\tau, \alpha^\tau, b_i^\tau|\tilde{\mu}, \sigma_i).$$

Definition 5. A mediated strategy $\tilde{\mu}$ is called a *communication equilibrium*, or just an *equilibrium*, if every unilateral deviation from $\tilde{\mu}$ is unprofitable:

$$U_i(\tilde{\mu}) \geq U_i(\tilde{\mu}|\sigma_i) \quad \forall (i, \sigma_i).$$

4.1 Public versus Private Equilibrium

Definition 5 is close to Myerson's definition, where the mediator uses a canonical communication system consisting of secret recommendations of what actions to take, without public announcements. I added public announcements to distinguish public and private communication equilibria, as the former enjoy a tractable recursive structure. Myerson's definition is a special case of Definition 5 that I label "private."

Formally, a mediated strategy $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2)$ is *private* if $\mathcal{A}$ is a singleton or $(\tilde{\mu}_1, \tilde{\mu}_2)$ is independent of past realizations of α. A private mediated strategy that is also a communication equilibrium is called a *private communication equilibrium*, or just a *private equilibrium*.

A mediated strategy $\tilde{\mu}$ induces a private mediated strategy $\tilde{\nu}$ by integrating out $\mathcal{A}$:

$$\tilde{\nu}(a_{\tau+1}|a^\tau, \omega^\tau) = \frac{\sum_{\alpha^\tau} \tilde{\mu}_1(a_{\tau+1}|a^\tau, \omega^\tau, \alpha^\tau) \prod_s \tilde{\mu}_1(a_s|a^{s-1}, \omega^{s-1}, \alpha^{s-1}) \tilde{\mu}_2(\alpha_s|a^s, \omega^s, \alpha^{s-1})}{\sum_{\alpha^\tau} \prod_s \tilde{\mu}_1(a_s|a^{s-1}, \omega^{s-1}, \alpha^{s-1}) \tilde{\mu}_2(\alpha_s|a^s, \omega^s, \alpha^{s-1})}.$$

It follows that $\Pr(a^\tau, \omega^\tau|\tilde{\nu}) = \Pr(a^\tau, \omega^\tau|\tilde{\mu})$ for every (a^τ, ω^τ), therefore $U_i(\tilde{\nu}) = U_i(\tilde{\mu})$.

Lemma 5. *If $\tilde{\mu}$ is an equilibrium then $\tilde{\nu}$ is a private equilibrium.*

Private equilibria are, in some sense, most general, as they enjoy the fewest incentive constraints. However, they are difficult to analyze. This motivates the study of public equilibria. Let $H_p = \bigcup_\tau H_p^\tau$ be the set of partial *public histories*, where $H_p^\tau = (\Omega \times \mathcal{A})^{\tau-1}$ collects the public information available to players up to period τ. A *public deviation* from $\tilde{\mu}$

is a map $\overline{\sigma}_i : H_p \to M(A_i)$. In words, a public deviation only depends on public information. Specifically, a player's mapping from current recommendations to mixed strategies only depends on public information. Say that a mediated strategy $\tilde{\mu}$ *discourages public deviations* if every public deviation is unprofitable, i.e., $U_i(\tilde{\mu}) \geq U_i(\tilde{\mu}|\overline{\sigma}_i)$ for all $(i, \overline{\sigma}_i)$.

Definition 6. A mediated strategy $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2)$ is *public* if it publicly announces all previous recommendations: $\mathcal{A} = A$ and $\tilde{\mu}_2(a_\tau|a^\tau, \omega^\tau, \alpha^{\tau-1}) = 1$ for all $(\tau, a^\tau, \omega^\tau, \alpha^{\tau-1})$. Therefore, without loss we will use $\tilde{\mu}$ and $\tilde{\mu}_1$ interchangeably in this case. A public mediated strategy that is also an equilibrium is a *public equilibrium.*[13]

Proposition 2. *A public mediated strategy that discourages public deviations is an equilibrium, hence a public equilibrium.*

Proposition 2 allows us to establish that public equilibria enjoy a tractably recursive structure, as usual. Intuitively, a player's past deviations do not affect his beliefs about opponents' future behavior. Formally, if $\tilde{\mu}$ is a public mediated strategy and $h_p^\tau \in H_p^\tau$ a partial public history then we may rewrite a player i's payoffs as follows:

$$v_i(h_p^t) = (1-\delta)\sum_{a_\tau} u_i(a_\tau)\tilde{\mu}(a_\tau|h_p^\tau) + \delta \sum_{(a_\tau,\omega_\tau)} v_i(a_\tau, \omega_\tau, h_p^t)\Pr(\omega_\tau|a_\tau)\tilde{\mu}(a_\tau|h_p^\tau).$$

The public mediated strategy $\tilde{\mu}$ is therefore a *public equilibrium* if for every player i, public history h_p^τ and one-shot deviation $\sigma_{it} \in M(A_i)$,

$$v_i(h_p^\tau) \geq \sum_{(a_\tau, b_{i\tau}, \omega_\tau)} [(1-\delta)u_i(b_{i\tau}, a_{-i\tau}) + \delta v_i(a_\tau, \omega_\tau, h_p^\tau)]\Pr(\omega_\tau|b_{i\tau}, a_{-i\tau})\sigma_{i\tau}(b_{i\tau}|a_{i\tau})\tilde{\mu}(a_\tau|h_p^\tau).$$

4.2 T-Public Equilibrium

Let us now define T-public communication equilibrium. Given a block length $T \in \mathbb{N}$, the mediated strategy $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2)$ is called *T-public* if $\mathcal{A} = A^T \cup \{0\}$ and

$$\tilde{\mu}_2(\alpha_\tau|a^\tau, \omega^\tau, \alpha^{\tau-1}) = 1 \text{ if } \begin{cases} \alpha_\tau = 0 \text{ and } \tau \neq kT \text{ for some } k \in \mathbb{N}, \text{ and} \\ \alpha_\tau = a_{\tau-T+1}^\tau \text{ and } \tau = kT \text{ for some } k \in \mathbb{N}, \end{cases}$$

where $a_{\tau-T+1}^\tau = (a_{\tau-T+1}, \ldots, a_\tau)$ lists the recommendation profiles in the most recent T-period block. In words, a T-public mediated strategy publicly announces all of the mediator's recommendations every T periods. Again, to economize on notation, I identify $\tilde{\mu}$ with $\tilde{\mu}_1$ and understand $\tilde{\mu}_2$ implicitly as just defined.

[13]This notion of public equilibrium is comparable to the recursive communication equilibrium introduced by Tomala (2009) and Rahman and Obara (2010) for games in discrete time.

A T-*public equilibrium* is a T-public mediated strategy that is also an equilibrium. An analogue of Proposition 2 applies to T-public equilibria as follows. The set of T-public histories is denoted by $H_p^T = \bigcup_\tau H_p^{T\lfloor \tau/T \rfloor}$, where $H_p^{Tk} = (\Omega^T \times A^T)^k$ is the set of partial public histories in any period τ such that $k = \lfloor \tau/T \rfloor$. Let $\mathcal{M}_i^T = \prod_{\tau=1}^T \{H_i^\tau \to \Delta(A_i)\}$ be the set of private-history-contingent mixed strategies within a T-period block. They represent a player's plan to privately deviate along a block. A T-*public deviation* from $\tilde{\mu}$ is a map $\sigma_i : H_p^T \to \mathcal{M}_i^T$. Thus, σ_i describes how a player plans to privately deviate at the beginning of each T-period block.

By Proposition 2 applied to T-period blocks, the T-public mediated strategy $\tilde{\mu}$ is a T-public equilibrium if it discourages T-public deviations, that is, for every player i, stage $k \in \mathbb{N}$, public kT-period public history $h_p^{Tk} \in H_p^{Tk}$ and T-public deviation σ_i,

$$v_i(h_p^{Tk}) \geq \sum_{h_{iTk}^{T(k+1)}} [(1-\delta) \sum_{\tau=Tk+1}^{T(k+1)} \delta^{\tau-1} u_i(b_{i\tau}, a_{-i\tau}) + \delta^T v_i(a^T, \omega^T, h_p^{Tk})] \times$$
$$\prod_{\tau=Tk+1}^{T(k+1)} \Pr(\omega_\tau | b_{i\tau}, a_{-i\tau}) \sigma_{i\tau}(b_{i\tau} | a_{i\tau}, h_{iTk}^\tau, h_p^{Tk}) \tilde{\mu}(a_\tau | h_p^\tau),$$

where $h_{iTk}^\tau \in H_i^\tau$ is a partial private history for player i in the kth T-period block.

4.3 Public versus Private Monitoring

The equilibrium concepts above apply equally to games with private monitoring modulo the following minor changes. Let $R_i = \{\rho_i : \Omega_i \to \Omega_i\}$ be the set of reporting strategies, where $\rho_i(\omega_i)$ is player i's report after observing ω_i. Now, $H_i^\tau = (A_i \times \Omega_i \times A_i \times \Omega_i \times \mathcal{A})^{\tau-1}$, which includes private observations and reports, and a deviation is any $\hat{\sigma}_i$ such that $\hat{\sigma}_i(a_i', \rho_i | a_i, h_i^\tau)$ is the probability that i deviates to (a_i', ρ_i) when asked to play a_i at private history h_i^τ.

The probability of $(a^\tau, \omega^\tau, \alpha^\tau, a_i^\tau{}', \omega_i^\tau{}')$ if i plays $\hat{\sigma}_i$ equals

$$\Pr(a^\tau, \omega^\tau, \alpha^\tau, a_i^\tau{}', \omega_i^\tau{}' | \tilde{\mu}, \hat{\sigma}_i) =$$
$$\Pr(a^\tau, \omega^\tau, \alpha^\tau | \tilde{\mu}) \prod_{s=1}^{\tau} \sum_{\{\rho_{is} : \rho_{is}(\omega_{is}') = \omega_{is}\}} \hat{\sigma}_i(a_{is}', \rho_{is} | a_i^s, \omega_i^{s-1}{}', a_i^{s-1}{}', \omega^{s-1}, \alpha^{s-1}) \frac{\Pr(\omega_{is}', \omega_{-is} | a_{is}', a_{-is})}{\Pr(\omega_s | a_s)}.$$

Therefore, the utility to player i from a unilateral deviation $\hat{\sigma}_i$ can be written as

$$U_i(\tilde{\mu} | \hat{\sigma}_i) = (1-\delta) \sum_{(\tau, a^\tau, a_i^\tau{}', \omega^\tau, \omega_i^\tau{}', \alpha^\tau)} \delta^{\tau-1} u_i(a_{i\tau}', a_{-i\tau}) \Pr(a^\tau, \omega^\tau, \alpha^\tau, a_i^\tau{}', \omega_i^\tau{}' | \tilde{\mu}, \hat{\sigma}_i).$$

After amending a player's private history, deviations, and their consequences, the definitions of equilibrium as well as the propositions above—including the distinction between private, public and T-public—can now be applied directly to the case of private monitoring.

5 Folk Theorems

Below, I begin in Section 5.1 an introductory discussion of how the example of Section 2 applies to general repeated games, which motivates an additional restriction on payoffs. In Section 5.2, I make such restriction, state the Folk Theorem and provide a brief outline of the proof, which can be found in Appendix A. Then I offer an immediate extension. In Section 5.3, I remove the additional restriction on payoffs but impose a slightly stronger condition on drifts that permits joint punishments without forgoing the need to burn value. In Section 6, I discuss other extensions of these results in several directions.

5.1 Summary and Intuition

Sometimes, as in Lemma 1, it is possible to give players of a game appropriate incentives without the need to burn value, even if it requires some amount of sophistication in terms of monitoring and communication. Other times, as in Lemma 2, value-burning simply cannot be avoided. The main results of the paper show how value-burning can be accomplished with frequent actions, in contrast with Sannikov and Skrzypacz (2010). The first result relies on individual punishments and rewards, the second on joint punishments.

On the one hand, the noise-to signal ratio from Brownian information explodes over an infinitesimal interval of time, so any value-burning becomes infeasible in public strategies. However, by avoiding punishment during some periods, it is possible to temper the noise-to-signal ratio and restore the feasibility of value-burning. Just when punishment is avoided must be kept secret from players to maintain incentives. Eventually, though, evaluation of punishments must arrive. For simplicity, I fix a length of calendar time $c > 0$ during which punishment periods are kept secret, with evaluations at the end of every block of length c. With respect to these c-length blocks of time, the game behaves similarly to a repeated game with mediated communication in discrete time. As such, I extend and apply the broad approach of Fudenberg et al. (1994) to this c-block repeated game. However, instead of using efficient "budget-balanced" incentives, which avoid value-burning, I use individual punishments and rewards to establish my first main result, Theorem 2.

Intuitively, the result is established as follows. Just as in Fudenberg et al. (1994), consider a smooth set W of candidate equilibrium payoffs in the interior of the some subset (to be explained later) of the feasible, individually rational set. I show that every payoff profile in W is locally self-decomposable under conditional identifiability. To this end, for any boundary point of W, player i is given incentives via individual punishments if $\lambda_i \geq 0$ and individual rewards if $\lambda_i < 0$, where $\lambda \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ is an outward normal vector to the boundary of W through the given point. Thus, $\lambda \geq 0$ for an efficient point of W.

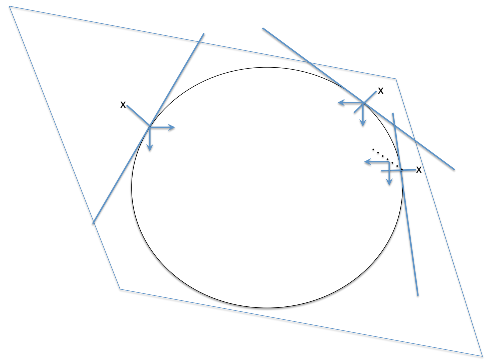

Figure 4: Local self-decomposability in every direction

In order to be able to use individual punishments and rewards, I rely on conditional (unconditional) identifiability for the case of public (private) monitoring. This is neither enough to avoid value-burning, as in Lemma 2, nor to induce joint punishments (where every player is punished simultaneously), as I discuss later. From Figure 4, it is clear that some payoff profiles must necessarily be sustained with individual rewards. There is an inherent difference between punishments and rewards, though. When incentives are successfully provided with rewards, a profitable deviation must decrease the probability of reward, so in equilibrium the reward must compensate a player from every deviation, since he can always choose to deviate. As a result, any payoff sustained with rewards must be exceed the payoff of the best deviation from an action profile sustaining it. Rewards are used to sustain low payoffs of a player, so this implication delivers an additional restriction on payoffs. If the payoff of a deviation from some action profile is too large then it may not be possible to sustain even efficient outcomes. This point was made before by Compte (1998, Theorem 2) in a different context. Similarly, I make and discuss a corresponding restriction on the set of sustainable payoffs to establish Theorem 2.

The first extension of Theorem 2 does away with this additional restriction on payoffs without avoiding value-burning, but makes an additional restriction on the information structure to allow for joint punishments. In Theorem 2, identifiability is used to ensure that for each player, a scoring rule can be constructed as in Lemma 2 to discourage deviations and remain secret in equilibrium. The additional restriction consists of being able to construct a single scoring rule with this property for every player. In terms of detecting deviations, this amounts to there not being two players whose joint deviations exactly cancel each others'—that is, their convolution is identical to no deviation at all. This assumption still precludes the use of budget-balanced incentives, since it is still perfectly possible for the identity of a unilateral deviator to be undetectable after a deviation. Hence, it is still the case that value-burning cannot generally be avoided. Nevertheless, it is still possible to temper the cost of punishments in equilibrium with infrequent coordination.

All these results are established together for both public and private monitoring. As the identifiability conditions of Section 3 for the cases of public and private monitoring suggest, the informational demands in the case of private monitoring are substantially weaker than those for public monitoring. This is because it is now easier to keep scoring rules secret. To see this, notice that over the course of a block of time a player is uninformed of others' actual private observations. As a result the mediator can use other players' observations to construct a secret scoring rule.

Generally, the requirement on signals imposed by identifiability becomes more stringent with more signals that a player observes, but less stringent with more signals that the player does not observe. In Section 6, I discuss this further and establish conditions for identifiability to hold generically. With private monitoring, I also show in Section 6 that UI is comparable to individual full rank, a condition that can be seen as approximately necessary for a Folk Theorem.

5.2 Individual Punishments and Rewards

Given a correlated strategy $\mu \in \Delta(A)$, recall that

$$u_i^d(\mu) = \max_{\sigma_i} \sum_{(a,b_i)} u_i(b_i, a_{-i})\mu(a)\sigma_i(b_i|a_i).$$

For any subset of players J and any $\varepsilon_1, \varepsilon_2 > 0$, let

$$\underline{V}_J^{\varepsilon_1,\varepsilon_2} = \{v \in U : \exists \mu \in \Delta^{\varepsilon_1}(A) \text{ s.t. } v_i \geq u_i^d(\mu) + \varepsilon_2 \ \forall i \notin J,\ v_i \leq u_i(\mu) - \varepsilon_2 \ \forall i \in J\},$$

where $\Delta^\varepsilon(A) = \{\mu \in \Delta(A) : \mu(a) \geq \varepsilon \ \forall a\}$. Finally, let $V^{**} = V_* \cap \operatorname{int} U_0$, where

$$V_*^{\varepsilon_1,\varepsilon_2} = \bigcap_{J \subset I} \underline{V}_J^{\varepsilon_1,\varepsilon_2} \qquad \text{and} \qquad V_* = \bigcup_{\varepsilon_1,\varepsilon_2>0} V_*^{\varepsilon_1,\varepsilon_2}.$$

The set V^{**} corresponds approximately to the set W^{**} of Compte (1998), but it is easy to see that $W^{**} \subset V^{**}$. This is because Compte (1998) only allows players outside of some set J to deviate from pure strategy profiles, as opposed to the correlated strategies with full support above. V^{**} is used for the Folk Theorem below. For example, it is easy to see that $V^{**} = \operatorname{int} U_0$ in the Prisoners' Dilemma. Basically, V^{**} is the set of payoff profiles such that for every subset of players, there is a correlated strategy whose payoff exceeds the payoff for the players in that subset and, for the complementary set of players, their deviation payoff is exceeded by the payoffs in the set. According to Compte (1998, Proposition 3), the requirement that V^{**} includes a sequence of payoff profiles that converge to the Pareto efficient frontier is generic in the space of payoffs.

Now we can state the paper's main result, namely that an "equilibrium threats" Folk Theorem holds as players become unboundedly patient, even with frequent actions. Two properties of the result are worth emphasizing. First, players engage in mediated communication. Although actions are arbitrarily frequent, their public communication rounds occur in discrete time, thus, with respect to privacy of players' strategies, time is endogenously discrete. Secondly, asymptotic efficiency is achieved with value-burning. This is possible, unlike in Sannikov and Skrzypacz (2010), because burning of value is delayed and conditioned on richer information, which, intuitively, tempers the noise-to-signal ratio.

Theorem 2 (Individual Incentives)**.** *Under Assumptions 1, 2 and 3, for any payoff profile u in the interior of V^{**}, there exists $(\underline{r}, \underline{\Delta})$ such that for all $(r, \Delta t) \leq (\underline{r}, \underline{\Delta})$, a communication equilibrium of the repeated game with parameters $(r, \Delta t)$ exists whose payoff profile is u.*

By Theorem 2, any profile in $\text{int}V^{**}$ is an equilibrium payoff if players are sufficiently patient and interact frequently enough. Since $c = T\Delta t$, fixing Δt but increasing T and decreasing r substitutes for decreasing Δt, which implies a Folk Theorem in discrete time. The proof of Theorem 2 appears in Appendix A. I offer a brief outline below.

First, I establish that every smooth subset of V^{**} is locally self-decomposable with respect to T-public equilibria of some appropriate calendar length of time $c \approx T\Delta t$. I derive a uniform bound on c to self-decompose W. Next, since I give players individual punishments and rewards, local self-decomposability is straightforward for payoff vectors belonging to the boundary of W whose outward normal vectors are regular, that is, not coordinate vectors. This is because, for regular vectors, punishments and rewards from v lie inside W for sufficiently patient players. On the other hand, for coordinate vectors there exist players for whom individual punishments and rewards necessarily displace continuation values outside of W, violating self-decomposability. To correct this issue, I shift players' continuation payoffs inside of W by an amount proportional to r. Now, incentive-providing payments are linear in r. However, W is smooth, so the scope of feasible continuation payoffs is proportional to $\sqrt{r}$. Following Theorem 4.1 of Fudenberg et al. (1994), there exists $r > 0$ sufficiently small that the incentives provided above lie inside W. Finally, I show that these incentive schemes continue to work as Δt and r are lowered, for fixed c.

Assumptions 1 and 2 are standard. Assumption 3 is used to construct scoring rules with the following key properties, as in Lemma 2:

1. Belief stability: (a) in equilibrium, players learn nothing about their own score, and (b) a deviation can only lower the drift in a player's score.

2. Implementability: it is possible to discourage every one-step deviation.

Formally, a simple duality argument yields a scoring rule as follows.

Lemma 6. *CI implies that for every completely mixed correlated strategy μ and every player i there exist $\xi_i : A \times \Omega \to [0,1]$ and $\gamma_i > 0$ such that*

$$\pi_i \sum_{a_{-i}} \Pr(\omega|a_{-i}, b_i)\mu(a) \leq \sum_{a_{-i}} \xi_i(a,\omega) \Pr(\omega|a_{-i}, b_i)\mu(a) \quad \forall (i, a_i, b_i, \omega),$$

where $\pi_i = \sum_{(a,\omega)} \xi_i(a,\omega) \Pr(\omega|a)\mu(a)$ for every player i, and

$$\gamma_i \sum_{a_{-i}} \Delta u_i(a, b_i)\mu(a) \leq \sum_{(a_{-i},\omega)} \xi_i(a,\omega)\Delta \Pr(\omega|a, b_i)\mu(a) \quad \forall (i, a_i, b_i),$$

where $\Delta u_i(a, b_i) = u_i(a_{-i}, b_i) - u_i(a)$ and $\Delta \Pr(\omega|a, b_i) = \Pr(\omega|a_{-i}, b_i) - \Pr(\omega|a)$. Such a function ξ will be called a proper scoring rule.

This result is the content of Proposition 6, and is proved there. The function ξ_i and number γ_i are used to construct a scoring rule as in Lemma 2 (γ_i corresponds to $1/w$ in the example) to give players incentives over a block of given length c. Appropriate cutoffs are also calculated to discourage deviations by not punishing too seldom and so that the cost of providing incentives can be made economical by not punishing too often.

The difference above between public and private monitoring is simply to do with the fact that in the latter (i) players can misreport their observations and must be discouraged from doing so, and (ii) they do not observe others' observations so cannot condition their behavior on them. With private monitoring, a scoring rule is constructed almost identically to the lemma above except for these minor caveats, as follows.

Corollary 1. *With private monitoring, UI implies that for some (ξ, γ),*

$$\pi_i \sum_{a_{-i}} \Pr(\omega_i', \omega_{-i}|a_{-i}, a_i')\mu(a) \leq \sum_{a_{-i}} \xi_i(a,\omega) \Pr(\omega_i', \omega_{-i}|a_{-i}, a_i', \rho_i)\mu(a) \;\; \forall (i, a_i, a_i', \omega_i, \omega_i'), \tag{5}$$

where $\pi_i = \sum_{(a,\omega)} \xi_i(a,\omega) \Pr(\omega|a)\mu(a)$ for every player i, and

$$\gamma_i \sum_{a_{-i}} \Delta u_i(a, a_i')\mu(a) \leq \sum_{(a_{-i},\omega)} \xi_i(a,\omega)\Delta \Pr(\omega|a, a_i', \rho_i)\mu(a) \quad \forall (i, a_i, a_i', \rho_i), \tag{6}$$

where $\rho_i : \Omega_i \to \Omega_i$ is a reporting strategy (that is, $\rho_i(\omega_i)$ is reported if ω_i is observed), $\Delta \Pr(\omega|a, a_i', \rho_i) = \Pr(\omega|a_{-i}, a_i', \rho_i) - \Pr(\omega|a)$ and

$$\Pr(\omega|a_{-i}, a_i', \rho_i) = \sum_{\omega_i' \in \rho_i^{-1}(\omega_i)} \Pr(\omega_{-i}, \omega_i'|a_{-i}, a_i')$$

is the probability that ω is reported when everyone is honest and plays their part of a except for player i, who plays a_i' instead and reports according to the reporting strategy ρ_i. Again, such a function ξ is called a proper scoring rule.

5.3 Joint Punishments

The previous restriction that payoffs lie inside V^{**} can be avoided altogether by slightly strengthening the information structure to allow for joint punishments. In this case, players are jointly punished with positive probability of static equilibrium henceforth.

Definition 7. Say that the matrix Pr exhibits *joint conditional identifiability* (JCI) if, whenever $y_i \in \mathbb{R}_+^{A_i \times A_i \times \Omega}$ for every player i,

$$\sum_{(i,a_i,b_i,\omega)} y_i(a_i, b_i, \omega) \Pr(\omega|a_{-i}, b_i) = \Pr(\omega|a) \;\; \forall(a, \omega) \quad \Rightarrow \quad y_i(a_i, b_i, \omega) = 0 \;\text{ for all }\; a_i \neq b_i,$$

and *joint unconditional identifiability* (JUI) if, whenever $y_i \in \mathbb{R}_+^{A_i \times A_i \times \Omega_i \times \Omega_i}$ for every i,

$$\sum_{(i,a_i,\omega_i,a_i',\omega_i')} y_i(a_i, \omega_i, a_i', \omega_i') \Pr(\omega_i', \omega_{-i}|a_{-i}, a_i') = \Pr(\omega|a) \;\; \forall(a, \omega) \quad \Rightarrow$$

$$y_i(a_i, a_i', \omega_i, \omega_i') = 0 \;\text{ for all }\; (a_i, \omega_i) \neq (a_i', \omega_i').$$

The drift x exhibits *joint conditional identifiability* (JCI-x) if, when $y_i \in \mathbb{R}_+^{A_i \times A_i}$ for all i,

$$\sum_{(i,a_i,b_i)} y_i(a_i, b_i)[x(a_{-i}, b_i) - x(a)] = 0 \;\; \forall a \in A \quad \Rightarrow \quad y_i(a_i, b_i) = 0 \;\text{ for all }\; a_i \neq b_i.$$

In contrast with the previous version of "individual" identifiability, joint identifiability requires that the statistical consequences of every deviation profile are not diametrically opposed for any two disjoint subsets of players. This way, when punishing everyone, these two sets of players cannot deviate in a way that precludes joint punishment. This is not inconsistent with obedient agents being unidentifiable after a deviation, see Example 5.

Example 5. Consider the repeated Cournot oligopoly with two firms. Each firm i can produce any amount $q_i \in \mathbb{R}$, not necessarily positive. The market price at time t equals $p_t = P(q_{1t} + q_{2t}) + \varepsilon_t$, where P is an inverse demand curve such that $P'(q) < 0$ for all q and $\varepsilon_t \sim N(0, \sigma^2/\Delta t)$ is a random shock. Any y such that $y_1(q_1, q_1 + h) = q_2(q_2, q_2 - h) > 0$ for some $h \neq 0$ (and 0 otherwise) satisfies the antecedent above, so joint identifiability fails.

It is easy to see, as in Lemma 6, that joint identifiability guarantees the existence of a common scoring rule $\xi : A \times \Omega \to \mathbb{R}$ with which to incentivize every player simultaneously. Closed joint identifiability ensures that such scoring rules converge meaningfully as $\Delta t \to 0$, which is necessary for the Folk Theorem below. To define it, given a profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ such that $\sigma_i : A_i \to \Delta(A_i)$ is a recommendation-contingent deviation plan for every i, let $\mathrm{Lr}(\omega|a, \sigma) = \Pr(\omega|a)/\sum_i \Pr(\omega|a_{-i}, \sigma_i)$ and $\Delta\mathrm{Lr}(\omega|\sigma) = \max_{(a,b)} \mathrm{Lr}(\omega|a, \sigma) - \mathrm{Lr}(\omega|b, \sigma)$. Say that Pr exhibits *closed joint conditional identifiability* ($\overline{\text{JCI}}$) if for every deviation profile σ

such that $\sigma_i(a_i|a_i) < 1$ for some (i, a_i), $\limsup_{\Delta t \to 0} \Delta \mathrm{Lr}(\omega|\sigma)/\sqrt{\Delta t} > 0$. When monitoring is private and $\tilde{\sigma}$ is a profile such that $\tilde{\sigma}_i : A_i \times \Omega_i \to \Delta(A_i \times \Omega_i)$ is a recommendation- and signal-contingent deviation plan for each i, let $\mathrm{Lr}(\omega|a, \tilde{\sigma}) = \Pr(\omega|a)/\sum_i \Pr(\omega_{-i}|a_{-i}, \tilde{\sigma}_i)$ and $\Delta \mathrm{Lr}(\omega|\tilde{\sigma}) = \max_{(a,\omega,a',\omega')} \mathrm{Lr}(\omega|a, \tilde{\sigma}) - \mathrm{Lr}(\omega'|a', \tilde{\sigma})$. Say that Pr exhibits *closed joint unconditional identifiability* ($\overline{\mathrm{JUI}}$) if for every $\tilde{\sigma}$ such that $\tilde{\sigma}_i(a_i, \omega_i|a_i, \omega_i) < 1$ for some (i, a_i, ω_i), it is the case that $\limsup_{\Delta t \to 0} \Delta \mathrm{Lr}(\omega|\tilde{\sigma})/\sqrt{\Delta t} > 0$.

Definition 8. *Closed joint identifiability* means $\overline{\mathrm{JCI}}$ ($\overline{\mathrm{JUI}}$) if monitoring is public (private).

For Lemma 2, every player was given incentives with an independent, latent score. Joint identifiability means that a single latent score can simultaneously provide incentives for every player, as I show next. The proof is very close to that of Lemma 6, so omitted.

Lemma 7. *If x exhibits joint conditional identifiability then for every completely mixed correlated strategy μ there exist $\xi : A \times \Omega \to [0, 1]$ and $\gamma > 0$ such that*

$$\pi \sum_{a_{-i}} \Pr(\omega|a_{-i}, b_i)\mu(a) \leq \sum_{a_{-i}} \xi(a, \omega) \Pr(\omega|a_{-i}, b_i)\mu(a) \quad \forall(i, a_i, b_i, \omega),$$

where $\pi = \sum_{(i,a,\omega)} \xi_i(a, \omega) \Pr(\omega|a)\mu(a)$ and

$$\gamma \sum_{a_{-i}} \Delta u_i(a, b_i)\mu(a) \leq \sum_{(a_{-i},\omega)} \xi(a, \omega)\omega\Delta x(a, b_i)\mu(a) \quad \forall(i, a_i, b_i).$$

A Folk Theorem follows from Lemma 7: players are jointly threatened with positive probability of a static equilibrium henceforth unless their joint score is sufficiently high.

Theorem 3 (Joint Punishments)**.** *Under Assumptions 1 and 2, closed joint identifiability implies that for any payoff profile u in the interior of U_0, there exists $(\underline{r}, \underline{\Delta})$ such that for all $(r, \Delta t) \leq (\underline{r}, \underline{\Delta})$, a communication equilibrium of the repeated game with parameters $(r, \Delta t)$ exists whose payoff profile is u.*

To illustrate just how joint punishments might work, recall the Prisoners' Dilemma of Section 2. I will construct a joint scoring rule assuming $x_1 \neq \frac{1}{2}(x_0 + x_2)$, as in Lemma 2.

Example 6. In the Prisoners' Dilemma, assume that $x_0 + x_2 < 2x_1$ and recall the setting of Lemma 2. To construct a joint latent variable Y_t, let $Y_0 = 0$ and $Y_t = Y_{t-\Delta t} \pm \sqrt{\Delta t}$ follow a random walk with failure probability $\xi(a_t, \omega_t)$, where a_t is the profile of recommendations at time t and ω_t the realized public signal. This failure probability is defined as follows: $\xi(a, \omega) = \frac{1}{2}$ whenever $\omega = -1$. For $\omega = +1$, let $\xi(DD, +1) = 1$,

$$\xi(CC, +1) = \tfrac{1}{2}\left[1 + \left(\frac{\mu}{1-\mu}\right)^2 \frac{p_0}{p_2}\right] \quad \text{and} \quad \xi(CD, +1) = \xi(DC, +1) = \tfrac{1}{2}\left[1 - \frac{\mu}{1-\mu}\frac{p_0}{p_1}\right].$$

It can be verified easily that for any signal realization, the scoring rule has no drift in equilibrium. For any deviation, however, the drift of Y_t is lowered significantly. As in the proof sketch of Lemma 2, the failure probability after defecting when asked to cooperate is

$$\begin{aligned}
\Pr(F_t|C_i, D_i, +1) &= \frac{(1-\mu)p_1\xi(CC, +1) + \mu p_0 \xi(CD, +1)}{(1-\mu)p_1 + \mu p_0} \\
&= \tfrac{1}{2}\left[1 + \frac{\mu^2/(1-\mu)}{(1-\mu)p_1 + \mu p_0}\frac{p_0(p_1^2 - p_0p_2)}{p_1p_2}\right] \approx \tfrac{1}{2}\left[1 + \frac{\mu^2}{1-\mu}(2x_1 - (x_0 + x_2))\sqrt{\Delta t}\right],
\end{aligned}$$

where F_t stands for Y_t jumping down in period t. The failure drift associated with this failure probability is clearly $\hat{z} = \frac{\mu^2}{1-\mu}(2x_1 - (x_0 + x_2))$. In contrast with Lemma 2, we must also check the failure drift from cooperating when asked to defect:

$$\begin{aligned}
\pi^{**} &= \Pr(F_t|D_i, C_i, +1) = \frac{(1-\mu)p_2\xi(CD, +1) + \mu p_1 \xi(DD, +1)}{(1-\mu)p_2 + \mu p_1} \\
&= \tfrac{1}{2}\left[1 + \frac{\mu}{(1-\mu)p_1 + \mu p_0}\frac{p_1^2 - p_0p_2}{p_1}\right] \approx \tfrac{1}{2}\left[1 + \mu(2x_1 - (x_0 + x_2))\sqrt{\Delta t}\right],
\end{aligned}$$

Let $z^{**} = \mu(2x_1 - (x_0 + x_2))$ be the failure drift associated with π^{**}. Assume that μ is sufficiently small for z^{**} to be the maximal failure drift, i.e., $z^{**} > \hat{z}$. The prior failure drift, as in Lemma 2, is $z^* = \frac{1}{2}\mu(2x_1 - (x_0 + x_2))$.

Equilibrium is constructed as follows given Y. The mediator makes recommendations and generates Y secretly for a length of calendar time $c > 0$. At the end of the block, if $Y_c \geq Y^{**} = -z^{**}c$ then everyone's continuation values remain at the previous level v. If $Y_c < Y^{**}$ then players resort to DD forever with some probability α such that $\alpha v = w$ from the proof sketch of Lemma 2.

The only substantial difference comes in checking that incentive compatibility constraints are satisfied. I go over these constraints next intuitively, since their logic follows that of Lemma 2 relatively closely. If a player was asked to defect, cooperating raises the probability of punishment and brings no current gain, so is unprofitable. If a player was asked to cooperate in the very first period, the previous analysis produces a punishment w that discourages defection. Finally, since z^{**} was chosen instead of $\hat{z} < z^{**}$ to determine the test for the scoring rule Y, there is no private history such that the density of the score is any less than $\varphi(z^{**}\sqrt{c})$, so incentives are never exhausted. As a result, all incentive constraints are satisfied by the same logic as in the proof sketch of Lemma 2.

Notice that this construction works for any initial lifetime equilibrium payoff profile strictly above the payoff profile from the static equilibrium that was used to punish players after a failed joint test. An equilibrium-threats Folk Theorem therefore emerges directly from this observation, so after taking the limit $\Delta t \to 0$, for every payoff profile $v \gg 0$ there is a discount rate $r > 0$ at which v is an equilibrium payoff.

6 Discussion

Next I discuss the model's assumptions and possible extensions.

6.1 Social Incentives

Theorem 2 assumes that incentives can only be given individually. In fact, the detectability assumptions there are the weakest sufficient conditions consistent with this restriction. However, it is natural to ask when it is possible to give social incentives, with schemes that correlate continuation payoffs across players. For instance, in Example 1, it was convenient and intuitive to do so when $x_2 \geq x_1 > x_0$, even if sometimes conditional identifiability held. At the same time, we obtained a Folk Theorem under weak conditions there. I address this general problem below in the context of public monitoring, although the same logic applies with private monitoring, as with the rest of the paper.

Definition 9. Given a vector of welfare weights λ, a proper λ-*balanced* scoring rule consists of failure probabilities ξ and a payment scheme $\beta : I \times A \times \Omega \to \mathbb{R}$ such that

1. *Belief stability*:

$$0 \leq \sum_{a_{-i}} (\xi_i(a, \omega) - \pi) \Pr(\omega | a_{-i}, b_i) \mu(a) \quad \forall (i, a_i, b_i, \omega), \tag{7}$$

where $\pi = \sum_{(i,a,\omega)} \xi_i(a, \omega) \Pr(\omega|a)\mu(a)$ for every player i.

2. *Implementability*: There exists $\gamma \in \mathbb{R}_+$ that yields incentive compatibility, i.e.,

$$\gamma \sum_{a_{-i}} \Delta u_i(a, b_i)\mu(a) \leq \sum_{(a_{-i},\omega)} \tfrac{\lambda_i}{|\lambda_i|} \xi_i(a, \omega) + \beta_i(a, \omega)\omega \Delta x(a, b_i)\mu(a) \ \forall (i, a_i, b_i);\text{[14]} \tag{8}$$

3. *Budget balance*: The payment scheme β is welfare neutral, i.e.,

$$\sum_{i=1}^{n} \lambda_i \beta_i(a, \omega) = 0 \quad \forall (a, \omega).$$

A balanced scoring rule has two new properties: (i) it includes a payment scheme β, where payments accrue each period with probability one but are only paid out at the end of the block, and (ii) the scoring rule is oriented around given welfare weights: if $\lambda_i > 0$ then ξ_i defines a punishment scheme, whereas if $\lambda_i < 0$ then ξ_i defines a reward scheme. If $\lambda_i = 0$ then player i is not given incentives with a scoring rule, only with payments β_i. Finally, the payment scheme β is budget-balanced around the vector of weights λ.

[14] I interpret $\frac{\lambda_i}{|\lambda_i|} = 0$ when $\lambda_i = 0$.

A deviation $\sigma_i \in M(A_i)$ is *conditionally irreversible* if

$$\Pr(\omega|a_i, \cdot) = y_i(a_i, \sigma_i, \omega) \Pr(\omega|\sigma_i(a_i), \cdot) + \sum_{b_i} y_i(a_i, b_i, \omega) \Pr(\omega|b_i, \cdot)$$

for some $y \geq 0$ and (i, a_i, ω) implies that $y_i(a_i, \sigma_i, \omega) = 0$. It is clear from the proof of Proposition 6 that what is necessary for it to hold is that every profitable deviation is conditionally irreversible with respect to Pr. Regarding drifts, we obtain a similar condition for Proposition 7. For Proposition 10 we qualify utility gains to be non-negative.

We seek identifiability conditions that guarantee incentive compatibility. A profile of deviations σ is λ-*unattributable* (Rahman and Obara, 2010) if there is a vector $\eta \in \mathbb{R}^{A \times \Omega}$ such that $\Delta \Pr(\omega|a, \sigma_i) = \lambda_i \eta(a, \omega)$ for every player i. The profile σ is *conditionally irreversible* with respect to λ if $\frac{\lambda_i}{|\lambda_i|}\sigma_i$ is conditionally irreversible for each i. Finally, σ is *profitable* if the sum of unilateral deviation gains across individuals is positive:

$$\sum_{i=1}^{n} \Delta u_i(\sigma_i, \mu) > 0.$$

Proposition 3. *Let μ be a completely mixed correlated strategy and λ a regular*[15] *vector of welfare weights. If every profitable, λ-unattributable deviation profile is conditionally irreversible with respect to λ then a proper λ-balanced scoring rule exists.*

If a profile of deviations is unattributable then—statistically speaking—as far as the mediator knows anyone could have been the deviator. If $\eta = 0$ then the deviations are undetectable and there is nothing that the mediator can do to prevent them, but if $\eta \neq 0$ then in order to provide incentives the mediator cannot punish some and reward others simultaneously, which generally leads to value-burning. To apply the punishment and reward schemes of Section 7, it is enough that (i) every individually profitable deviation is detectable, and (ii) profitable, unattributable deviations are conditionally irreversible. I should mention that Proposition 3 also applies in the limiting case with conditional irreversibility in drifts rather than probabilities, as in Section 7, as the proof suggests.

From Proposition 3, a somewhat weaker Folk Theorem than Theorem 2 obtains. For regular λ, punishment and reward schemes are designed as in Section 7, except that now, at the end of a T-period block, if the mediator's history was (a^T, ω^T), then each player i is pays the amount $\sum_\tau \beta_i(a_\tau, \omega_\tau)$ in units of continuation value in period T. The utility to player i over a T-period block includes the transfers β_i:

$$(1 - \delta^T)U_i'(a^T, \omega^T) = (1 - \delta) \sum_{\tau=1}^{T} [\delta^{\tau-1} u_{i\tau}(a_\tau) + \delta^T \beta_i(a_\tau, \omega_\tau)],$$

[15] Regular means that at least two elements of λ are non-zero, thus λ is not a coordinate vector.

where $U_i'(a^T, \omega^T)$ is player i's utility net of payments β_i. Now suppose that λ is a coordinate vector. If $\lambda = -\mathbf{1}_i$ then we can decompose a boundary payoff with a correlated equilibrium, as usual, whereas if $\lambda = \mathbf{1}_i$ assume that there is an enforceable correlated strategy that maximizes i's expected payoff. In this case, i needs no incentives. If c is small relative to $1/r$, players' net utility gains from a deviation will be bounded, so the incentive schemes of Section 7 apply and we can obtain a Folk Theorem in the same way as for Theorem 2, except that the detectability assumptions are the weaker ones of Proposition 3.

The detectability condition of Proposition 3 reconciles the characterization of a Folk Theorem in the Prisoners' Dilemma of Example 1. There, it was necessary and sufficient that every profitable deviation was detectable. Indeed, since players only have two actions, a profitable deviation simply cannot be reversed, hence the result for two-by-two games.

6.2 Identifiability is Generic

In case of public monitoring conditional identifiability may seem stringent, especially as the number of signals increases. Yet I show below that CI and CI-x are both generic if every player has at most as many actions as everyone else put together: $|A_i| \leq |A_{-i}|$ for all i. This means that, in the Binomial model, the Folk Theorem above can be applied to a generic family of drifts. One justification for the Binomial model is that—as shown by Fudenberg and Levine (2007, 2009)—the kind of information structure used to reach the continuous time limit can have crucial consequences. In the Binomial model, it is impossible to obtain positive results (i.e., overcome the impossibility of value-burning) with their approach.

With more signals per period, it can be shown that in the prototypical case of normally distributed signals, CI-x also suffices for existence of a proper scoring rule. For more general signal structures, it may be more difficult to satisfy conditional identifiability.

Intuitively, the notion behind conditional identifiability seems stronger than individual full rank—a leading notion of detectability for unilateral deviations in the literature on repeated games. With public monitoring, individual full rank (IFR) is usually defined as $\Pr(\cdot|a_i,\cdot) \notin \text{span}\{\Pr(\cdot|b_i,\cdot) : b_i \neq a_i\}$ for every (i, a_i), although it has been relaxed to the convex hull rather than the linear hull in some papers. Intuitively, IFR states that every unilateral deviation is statistically detectable. CI requires that a deviation be detectable *conditional* on each public signal, whereas IFR makes this requirement unconditionally. With private monitoring, unconditional identifiability is comparable to IFR, which can be thought of as $\Pr(\omega_i, \cdot|a_i, \cdot) \notin \text{span}\{\Pr(\omega_i', \cdot|a_i', \cdot) : (a_i', \omega_i') \neq (a_i, \omega_i)\}$ for every (i, a_i, ω_i). This is obviously close to UI and generic if $|\Omega_i| > 1$ and $|A_i \times \Omega_i| \leq |A_{-i} \times \Omega_{-i}|$ for all i.

Proposition 4. *CI is generic if $|A_i| \leq |A_{-i}|$ for all i. In case of private monitoring, UI is also generic if (a) $|A_i \times \Omega_i| \leq |A_{-i} \times \Omega_{-i}|$ when $|\Omega_i| > 1$ and $|\Omega_{-i}| > 1$, (b) $|A_i| - 1 \leq |A_{-i}|(|\Omega_{-i}| - 1)$ when $|\Omega_i| = 1$, and (c) $|A_i|(|\Omega_i| - 1) \leq |A_{-i}| - 1$ when $|\Omega_{-i}| = 1$.*

6.3 Dispensing with the Mediator

To some extent, the mediator can be dispensed with easily. This requires resorting to mixed strategy profiles rather than correlated strategies. Also, instead of taking recommendations, players may report their *intended* action before actually playing it.[16] The scoring rules derived above for recommendations ensure that players follow through on any intentions they report. One way to induce the right intentions is to solicit them at the beginning of every period, charge players in units of continuation value for their reports in order to keep them indifferent over reports, and subject them to punishments and rewards as defined above so that actual behavior agrees with the profile of reported intentions. To illustrate, assume that every player will be given incentives via punishments. Let μ_i be a mixed strategy for player i and $\mu_{-i} = \prod_{j \neq i} \mu_j$ the product of others' mixed strategies. Let $u_{i*} = \min_{a_i} \sum_{a_{-i}} u_i(a)\mu_{-i}(a_{-i})$ and $u_{i*}(a_i) = \sum_{a_{-i}} u_i(a)\mu_{-i}(a_{-i}) - u_{i*}$. Given an initial block of length c, if player i reports the intention to play a_{it} at date $t \in [0, c]$, he will be charged $w_{it}(a_{it}) = (1 - e^{-r\Delta t})e^{-rt}u_{i*}(a_{it})$ units of continuation value during subsequent blocks. In addition, player i will face a punishment scheme as defined in Section 7.2, where the mediator's recommendations are replaced with players' intentions. Given any intended action, a punishment scheme that satisfies incentive compatibility as described in Section 7.2 immediately satisfies incentive compatibility in this setting. Given that every player follows their reported intentions, the probability of ultimate punishment is independent of the actual reported intentions. Therefore, if a player abides by his intentions, different reported intentions only affect a player's payoff through the direct charge $w_{it}(a_{it})$. But this is constructed to keep a player indifferent over possible reports in this case. Hence:

Proposition 5. *An incentive compatible punishment or reward scheme remains incentive compatible when recommendations are replaced with intentions as described above.*

However, intentions are subject to Bhaskar's (2000) critique. For a player to reveal his intentions, he must be indifferent over reports and, playing mixed strategies, must also be indifferent over the pure actions in their support, too. With a mediator, these ties can be broken easily and robustly. Moreover, for some games detectability requires perfectly correlated behavior by some players, but crucially kept secret from others. In such games, it is unlikely that a mediator can be dispensed with. See Rahman (2012a) for details.

Even if a mediator cannot be dispensed with, it may be possible to decentralize it with plain conversation, as argued by Forges (1990, 1986); she required 4 or more players, though. If players communicated through actions instead of reporting intentions as above, any information communicated by players may be subject to additional incentive constraints.

[16]Relatedly, Kandori (2003) suggested that players report what they played to each other. Rahman (2012a) shows that, generally, it is better for players to report intentions than actual behavior.

7 Incentive Compatibility

To prove the Folk Theorems above, I now construct punishments and rewards for the repeated game. For simplicity, I restrict this section to the Binomial model of Example 3, but the appendix shows how all these results extend to the general case with both public and private monitoring. From now on, fix a stage-game correlated strategy $\mu \in \Delta(A)$.

7.1 Scoring Rules

Definition 10. A *scoring rule* is a pair of functions $\xi, \zeta : I \times A \times \Omega \to [0,1]$ of failure and success probabilities, respectively, with $\zeta = 1 - \xi$. Call ξ and ζ *proper*[17] if

1. *Belief stability*: Failure is equally likely with obedience, more likely without it.

$$\sum_{a_{-i}} \Pr(\omega|a_{-i}, b_i)\mu(a)\pi_i \leq \sum_{a_{-i}} \xi_i(a,\omega) \Pr(\omega|a_{-i}, b_i)\mu(a) \quad \forall (i, a_i, b_i, \omega), \tag{9}$$

   where $\pi_i = \sum_{(a,\omega)} \xi_i(a,\omega) \Pr(\omega|a)\mu(a)$ for every player i.

2. *Implementability*: Some $\gamma_i > 0$ yields incentive compatibility for each player i.

$$\gamma_i \sum_{a_{-i}} \Delta u_i(a, b_i)\mu(a) \leq \sum_{(a_{-i},\omega)} \xi_i(a,\omega)\omega\Delta x(a, b_i)\mu(a) \quad \forall (i, a_i, b_i), \tag{10}$$

   where $\Delta u_i(a, b_i) = u_i(a_{-i}, b_i) - u_i(a)$ and $\Delta x(a, b_i) = x(a_{-i}, b_i) - x(a)$.

A proper scoring rule consists of failure probabilities ξ (with punishment $\frac{1}{2}\sqrt{\Delta t}/\gamma_i$) such that (i) the probability of failure is at least as great after a deviation than after obeying the mediator's recommendations, and (ii) obedience can be incentive compatible. (With private monitoring, (9) and (10) become (5) and (6).) By belief stability, whenever a player obeys a recommendation, his posterior beliefs about the probability of failure conditional on his recommendation and signal equal his prior π_i, which can be chosen arbitrarily.

Lemma 8. (i) *If ξ is a proper scoring rule then conditional failure probabilities are constant and equal to π on the equilibrium path:*

$$\sum_{a_{-i}} \xi_i(a,\omega) \Pr(\omega|a)\mu(a) = \pi_i \sum_{a_{-i}} \Pr(\omega|a)\mu(a) \qquad \forall (i, a_i, \omega).$$

(ii) *If a proper scoring rule exists then for any vector $\alpha \in (0,1)^n$ there exists another proper scoring rule with prior failure probability for player i equal to α_i.*

Proposition 6. *If* Pr *exhibits conditional identifiability then the set of proper scoring rules for μ is not empty as long as μ is a completely mixed correlated strategy.*

[17] The label "proper" is borrowed from statistics (see, e.g., Gneiting and Raftery, 2007). Formally, belief stability implies properness, as its constraints are ex post, so a better label might be "very proper."

Notation. If $\sigma_i \in M(A_i)$ is a stage-game deviation and $f : A \to \mathbb{R}$ depends on actions, *not* recommendations, write $f(a,\sigma_i) = \sum_{b_i} f(a_{-i},b_i)\sigma_i(b_i|a_i)$ for the convolution of f with σ_i, $\Delta f(a,\sigma_i) = f(a,\sigma_i) - f(a)$ for the effect of σ_i on f at a, and $\Delta f(\mu,\sigma_i) = \sum_a \Delta f(a)\mu(a)$, or simply $\Delta f(\sigma_i)$, for the a priori effect. (With private monitoring, given $\hat{\sigma}_i : A_i \to \Delta(A_i \times R_i)$, let $f(a,\omega,\hat{\sigma}_i) = \sum_{(a_i',\rho_i,\omega_i' \in \rho_i^{-1}(\omega_i))} f(a_i',a_{-i},\omega_i',\omega_i)\hat{\sigma}_i(a_i',\rho_i|a_i)$, etc.)

Fix a proper scoring rule ξ. Let $\pi_i(\sigma_i) = \sum_{(a,\omega)} \xi_i(a,\omega)\Pr(\omega|a,\sigma_i)\mu(a)$ be the *prior failure probability* from σ_i and $\Delta\pi_i(\sigma_i) = \pi_i(\sigma_i) - \pi_i$. Let $z_i(\sigma_i) = \Delta\pi_i(\sigma_i)/(\frac{1}{2}\sqrt{\Delta t})$ be the *prior failure drift* from σ_i. The *incentive cost* of σ_i from ξ_i at μ equals the ratio $C_i(\sigma_i) = \Delta u_i(\sigma_i)/z_i(\sigma_i)$, where I assume that $0/0 = 0$. $C_i(\sigma_i)$ is the ratio of utility changes to drift changes. The *conditional failure probability* from σ_i given (a_i,ω) is written as

$$\pi_i(\sigma_i|a_i,\omega) = \frac{\sum_{a_{-i}} \xi_i(a,\omega)\Pr(\omega|a,\sigma_i)\mu(a)}{\sum_{a_{-i}} \Pr(\omega|a,\sigma_i)\mu(a)}.$$

Similarly, $\Delta\pi_i(\sigma_i|a_i,\omega) = \pi_i(\sigma_i|a_i,\omega) - \pi_i$ and $z_i(\sigma_i|a_i,\omega) = \Delta\pi_i(\sigma_i|a_i,\omega)/(\frac{1}{2}\sqrt{\Delta t})$. With private monitoring, $\pi_i(\hat{\sigma}_i|a_i,\omega_i)$ and $z_i(\hat{\sigma}_i|a_i,\omega_i)$ are defined similarly.

Definition 11. A *cost-maximizing* deviation for player i is any $\sigma_i^* \in \arg\max_{\sigma_i} C_i(\sigma_i)$. Let $\Delta u_i^* = \Delta u_i(\sigma_i^*)$, $z_i^* = z_i(\sigma_i^*)$ and $C_i^* = C_i(\sigma_i^*)$. A *failure-maximizing* conditional deviation for player i given (a_i,ω) is any $\sigma_i^{**}(a_i,\omega) \in \arg\max_{\sigma_i} \pi_i(\sigma_i|a_i,\omega)$. Player i's *maximum conditional failure probability* equals $\pi_i^{**} = \max_{(a_i,\omega)} \pi_i(\sigma_i^{**}(a_i,\omega)|a_i,\omega)$, with conditional failure drift $z_i^{**} = \Delta\pi_i^{**}/(\frac{1}{2}\sqrt{\Delta t})$, where $\Delta\pi_i^{**} = \pi_i^{**} - \pi_i$.

By implementability (10), $0 \le C_i^* < \infty$, so, since $M(A_i)$ is compact, σ_i^* and σ_i^{**} exist. If $C_i^* = 0$ then i needs no incentives, so without loss $C_i^* > 0$. Hence, $\pi_i^{**} > \pi_i$. By the Maximum Theorem, $\arg\max C_i$ and $\arg\max z_i$ have a continuous selection.

Lemma 9. *Let ξ be a proper scoring rule and μ completely mixed. Without loss, σ_i^* and σ_i^{**} are continuous with respect to ξ and μ. So are Δu_i^*, z_i^*, C_i^*, π_i^{**} and z_i^{**}.*

Finally, consider convergence of scoring rules. If $\xi \to \overline{\xi}$ for some scoring rule $\overline{\xi}$, then, by Lemma 9, $\sigma_i^* \to \overline{\sigma}_i^*$ and $\sigma_i^{**} \to \overline{\sigma}_i^{**}$ for some $(\overline{\sigma}_i^*, \overline{\sigma}_i^{**})$. Hence, $(\Delta u_i^*, z_i^*, C_i^*)$ converges to some $(\Delta\overline{u}_i^*, \overline{z}_i^*, \overline{C}_i^*)$ and (π_i^{**}, z_i^{**}) to some $(\overline{\pi}_i^{**}, \overline{z}_i^{**})$. By Proposition 6, if Pr exhibits conditional identifiability for all small $\Delta t > 0$ then every completely mixed correlated strategy has a proper scoring rule ξ such that $C_i^* < \infty$. Since $0 \le \xi \le 1$, without loss ξ converges to some scoring rule $\overline{\xi}$ as $\Delta t \to 0$.

Proposition 7. *CI-x implies that there is a family of proper scoring rules ξ with $\xi \to \overline{\xi}$ as $\Delta t \to 0$, where $\overline{\xi}$ is a proper scoring rule at $\Delta t = 0$. Hence, $\overline{\pi}_i \in (0,1)$ and $\overline{C}_i^* < \infty$.*

With private monitoring, correspondingly similar definitions and results apply, although conditional deviations and failure probabilities may now depend on (a_i,ω_i) instead of (a_i,ω).

7.2 Punishment Schemes

Definition 12. A *punishment scheme* is a triple (μ, ξ, w) with ξ a proper scoring rule for $\mu \in \Delta(A)$ and $w \in \mathbb{R}^n_+$ a vector of punishments. It is implemented as follows.

1. Over a T-period block, the mediator secretly recommends players to play the action profile $a^T = (a_1, \ldots, a_T)$ with probability $\prod_\tau \mu(a_\tau)$. Players only observe their own recommendations throughout the T-period block and these recommendations are independent and identically distributed, generated by μ.

2. Player i's *score* for the block is determined by the following secret process.

   (a) For any history (a^T, ω^T), where a^T is the vector of recommendations by the mediator and ω^T is the vector of realized public signals, the mediator performs T independent Bernoulli trials, called *scoring trials*. The trial at time t has failure probability $\xi_i(a_t, \omega_t)$ if a_t was recommended and ω_t realized.

   (b) Finally, the score equals the number of successes in the T scoring trials.

3. Punishment for player i ensues if his score at the end of the block does not exceed $\tau_i^{**} = \lfloor (1 - \pi_i^{**})(T - 1) \rfloor$, where σ_i^{**} is a failure-maximizing conditional deviation and π_i^{**} is player i's maximum conditional failure probability from Definition 11.

4. Punishment to player i entails subtracting w_i from i's continuation value. Otherwise, "no punishment" entails no change to player i's continuation value.

The probability that i is punished if everyone obeys the mediator is given by

$$\Pi_{i0} = \sum_{\tau=0}^{\tau_i^{**}} \binom{T}{\tau} \pi_i^{T-\tau} (1 - \pi_i)^\tau,$$

where $\pi_i = \sum_{(a, b_i, \omega)} \xi_i(a, \omega) \Pr(\omega|a) \mu(a)$ is the equilibrium prior failure probability. Therefore, the average lifetime utility to player i is given by

$$v_i = (1 - \delta^T) u_i(\mu) + \delta^T [(1 - \Pi_{i0}) v_i + \Pi_{i0}(v_i - w_i)],$$

where $u_i(\mu) = \sum_a u_i(a) \mu(a)$. Rearranging yields

$$v_i = u_i(\mu) - \frac{\delta^T}{1 - \delta^T} \Pi_{i0} w_i.$$

Having defined a punishment scheme, let us argue its incentive compatibility. First, I find minimal punishments that discourage player i from deviating in the first period of a block. Then I show that this punishment discourages every deviation.

If player i plans to only deviate in the first period by playing $\sigma_i \in M(A_i)$ instead of obeying the mediator's recommendation, his utility gain is $(1-\delta)\Delta u_i(\sigma_i)$. On the other hand, the additional cost from deviating is given by the present value punishment times the change in punishment probability. Let $\Pi_{i1}(\sigma_i)$ be the punishment probability if player i only disobeys in the first period, and does so according to σ_i. Discouraging σ_i requires

$$(1-\delta)\Delta u_i(\sigma_i) \leq e^{-rc} w_i [\Pi_{i1}(\sigma_i) - \Pi_{i0}]. \tag{11}$$

From period 2 onwards, player i will obey the mediator, so failure probability equals π_i then. Hence, if $F_{iT}(\tau)$ stands for the CDF of a Binomial random variable with failure probability π_i and T trials then

$$\Pi_{i1}(\sigma_i) = \pi_i(\sigma_i) F_{iT-1}(\tau_i^{**}) + (1-\pi_i(\sigma_i)) F_{iT-1}(\tau_i^{**}-1).$$

Therefore, letting $f_{iT}(t)$ stand for the probability mass function obtained from F_i, it follows that $f_{iT-1}(\tau_i^{**}) = F_{iT-1}(\tau_i^{**}) - F_{iT-1}(\tau_i^{**}-1)$, so (11) can be rewritten as

$$(1-\delta)\Delta u_i(\sigma_i) \leq e^{-rc} w_i \Delta\pi_i(\sigma_i) f_{iT-1}(\tau_i^{**}).$$

Lemma 10. *The punishment scheme above discourages every first-period deviation if*

$$rce^{rc}\frac{\Delta u_i^*}{\Delta \pi_i^*} \leq w_i T \binom{T}{\tau_i^{**}} \pi_i^{T-\tau_i^{**}} (1-\pi_i)^{\tau_i^{**}}. \tag{12}$$

I will now argue that every deviation is discouraged by a punishment scheme satisfying (12). I will show that for any partial history of deviations and observations h_i^τ, every one-step deviation in period $\tau \leq T$ followed by obedience henceforth on the part of player i is unprofitable. This clearly renders every dynamic deviation unprofitable, since every such dynamic deviation must have a history after which its last one-step deviation takes place. Indeed, given h_i^τ and a one-step deviation $\sigma_i \in M(A_i)$ at time t, let $F_{iT}(\tau_i^{**}|h_i^\tau, \sigma_i)$ be the probability that player i is punished if he deviates according to σ_i and let $F_{iT}(\tau_i^{**}|h_i^\tau)$ be the probability that he is punished if he chooses not to deviate after h_i^τ.

The utility gain from σ_i given h_i^τ is $(1-\delta)\Delta u_i(\sigma_i)$. On the other hand, the deviation costs at least $e^{-rc} w_i [F_{iT}(\tau_i^{**}|h_i^\tau, \sigma_i) - F_{iT}(\tau_i^{**}|h_i^\tau)]$. Let $F_{iT-1}(\cdot|h_i^\tau)$ be the CDF of the number of failures given h_i^τ during all periods except τ, assuming a failure probability of π_i for all periods larger than τ. Letting $f_{iT-1}(\cdot|h_i^\tau)$ be the probability mass function induced by $F_{iT-1}(\cdot|h_i^\tau)$, it easily follows—as for discouraging the first-period deviation just before Lemma 10—that σ_i is discouraged if

$$(1-\delta)\Delta u_i(\sigma_i) \leq e^{-rc} w_i \Delta\pi_i(\sigma_i) f_{iT-1}(\tau_i^{**}|h_i^\tau).$$

By belief stability (9), no partial history of deviations and observations can decrease the probability of failure below π_i every period. Using this observation, the next result shows that $f_{iT-1}(\tau_i^{**}|h_i^\tau) \geq f_{iT-1}(\tau_i^{**})$, that is, the probability of τ_i^{**} successes is smaller whenever player i obeyed the mediator's recommendations. Together with Lemma 10, this implies that discouraging every single one-step deviation discourages the last one-step deviation associated with any arbitrarily complex dynamic deviation. Proceeding by induction, this ultimately discourages every dynamic deviation.

Lemma 11. *For every history h_i^τ, it is the case that $f_{iT-1}(\tau_i^{**}|h_i^\tau) \geq f_{iT-1}(\tau_i^{**})$, therefore every dynamic deviation is unprofitable whenever w_i satisfies (12).*

This result is one of the key observations used to prove the Folk Theorem. It says that the tightest incentive constraint is deviating in the first period only. The heart of the argument involves showing that while the gain from deviating is bounded by a linear function of how many times one deviates, the cost due to increased punishment is convex (see Figure 5). This is not generally enough, though. For instance, the information-delay construction of Abreu et al. (1991) is exponential, hence convex, but the probability of punishment is so low that incentives completely break down as $\Delta t \to 0$ for any fixed $r > 0$.

Let us now turn to convergence. I first take the limit as $\Delta t \to 0$ and then, afterwards, the limit as $r \to 0$. Consider a family ξ' of proper scoring rules indexed by $\Delta t \geq 0$. By Proposition 7, it is possible to construct such a family and satisfy $\xi' \to \overline{\xi}'$ as $\Delta t \to 0$, where $\overline{\xi}'$ is a proper scoring rule at $\Delta t = 0$. Just to ease notation eventually, apply an affine transformation (Lemma 8(ii)) to each scoring rule in the family to obtain a new family of scoring rules ξ such that $\pi_i = \frac{1}{2}$ for every $\Delta t \geq 0$. I will fix this family ξ for the remaining discussion of punishment schemes.

Proposition 8.

1. *For every $\varepsilon > 0$ there exists $\underline{\Delta} > 0$ such that if*

$$w_i \geq \frac{rce^{rc}\Delta\overline{u}_i^*}{(1-\varepsilon)\varphi(\overline{z}_i^{**}\sqrt{c})\overline{z}_i^*\sqrt{c}} \tag{13}$$

   then every deviation is unprofitable for all $\Delta t \in (0, \underline{\Delta})$.

2. *If w_i is chosen so that (13) holds with equality then lifetime utility satisfies*

$$v_i \to u_i(\mu) - \frac{rc\Delta\overline{u}_i^*}{1-e^{-rc}} \frac{1-\Phi(\overline{z}_i^{**}\sqrt{c})}{\varphi(\overline{z}_i^{**}\sqrt{c})\overline{z}_i^*\sqrt{c}} \quad \text{as} \quad \Delta t \to 0.$$

3. *Finally, letting $c \to \infty$ but $rc \to 0$ as $r \to 0$,*

$$\lim_{r\to 0}\lim_{\Delta t\to 0} v_i = u_i(\mu).$$

7.3 Reward Schemes

Definition 13. A *reward scheme* is a triple (μ, ζ, w) such that $\xi = 1 - \zeta$ is a proper scoring rule for μ and $w \in \mathbb{R}^n_+$ is a vector of rewards. It is implemented as follows.

1. Recommendations follow the same rule as for punishment schemes.
2. The *score* of the block is determined as follows.
   (a) For any history (a^T, ω^T), where a^T is the vector of recommendations by the mediator and ω^T is the vector of realized public signals, the mediator performs T independent Bernoulli trials, called *scoring trials*. The trial at time t has success probability $\zeta_i(a_t, \omega_t)$ if a_t was recommended and ω_t realized.
   (b) Finally, player i's score equals the number of failures in the T scoring trials.
3. Reward for i ensues if his score at the end of the block is less than or equal to $\tau_{i*} = \lfloor \pi_i(T-1) \rfloor - 1$, where $\pi_i = 1 - \sum_{(a,\omega)} \zeta_i(a, \omega) \Pr(\omega|a)\mu(a)$.
4. Reward to player i entails adding w_i to i's continuation value.

Reward schemes are only slightly different from punishment schemes. Of course, reward schemes involve the possibility of increasing continuation payoffs, rather than lowering them. Since a player could deviate every period, the reward w_i must compensate for at least these deviation gains. Another notable difference is that the cut-off proportion of successes that induces reward is roughly the prior reward probability assuming obedient behavior. This is in stark contrast with punishment schemes, where the cut-off is the prior punishment probability assuming disobedient behavior. This yields *concave* deviation costs, so the most profitable way to deviate is to do so (almost) every period.

The lifetime utility of player i is given by

$$v_i = u_i(\mu) + \frac{\delta^T}{1 - \delta^T} R_{i0} w_i,$$

where R_{i0}, the probability of reward given obedience, equals

$$R_{i0} = \sum_{\tau=0}^{\tau_{i*}} \binom{T}{\tau} \pi_i^\tau (1 - \pi_i)^{T-\tau}.$$

Let σ_i^T be a dynamic deviation for a T-period block. Discouraging σ_i^T entails

$$(1 - \delta) E \sum_{\tau=1}^{T} \delta^{\tau-1} \Delta u_i(\sigma_{i\tau}) \le \delta^T w_i E[R_{i0} - R_{iT}(\sigma_i^T)],$$

where R_{i0} and $R_{iT}(\sigma_i^T)$ are reward probabilities after obedience and σ_i^T respectively, and E, of course, computes expectations.

Let us find a useful bound on the reward w_i for incentive compatibility. For any terminal history of the block, h_i^T, let $R_{iT}(\sigma_i^T|h_i^T)$ be the reward probability from playing σ_i^T conditional on h_i^T. By belief stability (9), the failure probability $\pi_{i\tau}$ of each period's scoring trial must lie between π_i and π_i^{**}. Therefore, Hoeffding's (1956) inequality implies that

$$R_{iT}(\sigma_i^T|h_i^T) \leq \sum_{\tau=0}^{\tau_{i*}} \binom{T}{\tau} \hat{\pi}_i^\tau (1-\hat{\pi}_i)^{T-\tau} =: R_i(\hat{\pi}_i), \tag{14}$$

where $R_{iT}(\sigma_i^T|h_i^T)$ is the probability of being rewarded given h_i^T and $\hat{\pi}_i = \frac{1}{T}\sum_\tau \pi_{i\tau}$.

Lemma 12. *$R_i(\hat{\pi}_i)$ is strictly decreasing and strictly convex on $[\pi_i, \pi_i^{**}]$.*

Lemma 12 shows that—all else equal—player i prefers more variation in average failure probabilities. It also shows how reward schemes contrast punishment schemes. In Lemma 11, the choice of cut-off helped to determine that discouraging one deviation discouraged them all. Here, a different choice of cut-off yields the opposite conclusion: deviating only once is suboptimal is worse than deviating more often. Figure 5 (see Rahman, 2013, for details) illustrates this point as $\Delta t \to 0$. It shows how punishments are convex in the number of deviations, whereas rewards are concave. Hence, it is no longer the case for rewards that discouraging a first-period deviation discourages all deviations.

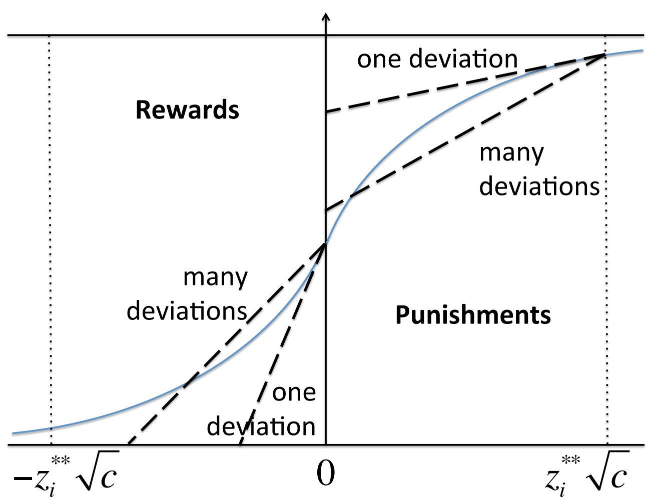

Figure 5: Concavity and convexity of punishments and rewards

Now, to bound w_i, let $E[\Delta\hat{u}_i(\sigma_i^T)] = \frac{1}{T}\sum_\tau \max\{E[\Delta u_i(\sigma_{i\tau})], 0\}$. The next result follows almost immediately from the incentive constraints above. A proof is therefore omitted.

Proposition 9. *The reward scheme above discourages every deviation if*

$$w_i \geq rce^{rc} \max_{\sigma_i^T} \frac{E[\Delta\hat{u}_i(\sigma_i^T)]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]}.$$

Let us now turn to convergence for reward schemes. Just as with punishment schemes, I first take the limit as $\Delta t \to 0$ and then the limit as $r \to 0$. Let ξ' be a family of proper scoring rules indexed by $\Delta t \geq 0$. Again, $\xi' \to \overline{\xi}'$ as $\Delta t \to 0$, where $\overline{\xi}'$ is a proper scoring rule at $\Delta t = 0$. Finally, apply an affine transformation (Lemma 8(ii)) to each scoring rule in the family to obtain a new family of scoring rules ξ such that $\pi_i = \frac{1}{2}$ for every $\Delta t \geq 0$. Fix this family ξ for the rest of this section. Given c and $T = \lfloor c/\Delta t \rfloor$, let

$$D_i = \sup_T \max_{\sigma_i^T} \frac{E[\Delta \hat{u}_i(\sigma_i^T)]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]}.$$

Proposition 10.

1. *A reward scheme with $w_i \geq rce^{rc} D_i$ discourages every deviation for all $\Delta t \in (0, c]$.*
2. *If $w_i = rce^{rc} D_i$ then player i's lifetime utility satisfies*

$$v_i \to u_i(\mu) + \frac{rc}{1 - e^{-rc}} \tfrac{1}{2} D_i \quad \text{as} \quad \Delta t \to 0.$$

3. *Finally, $D_i < \infty$, and letting $c \to \infty$ but $rc \to 0$ as $r \to 0$,*

$$\lim_{r \to 0} \lim_{\Delta t \to 0} v_i = u_i(\mu) + \Delta u_i^d,$$

where $\Delta u_i^d = \max_{\sigma_i} \Delta u_i(\sigma_i)$.

This result explains how incentives are provided with reward schemes in the long run, as players become patient. To allocate rewards more accurately, players can increase the length of their block and receive more signals with which to aggregate information. In the limit, reward schemes are efficient in the sense that players are compensated for their best deviation gains (which I already argued was unavoidable) and no more.

7.4 Self-Generation

In the appendix, I show how the punishments and rewards derived above translate into communication equilibria of the repeated game. Basically, I follow a T-augmentation of the algorithm by Fudenberg et al. (1994). As illustrated in Figure 4, consider a smooth subset of feasible payoff profiles W. Every payoff profile on the boundary is associated with an outward normal direction λ that describes which player i is given incentives with punishments ($\lambda_i > 0$) and which with rewards ($\lambda_i < 0$). Using as a stage game the T-period repetition of the original stage game, I use these punishments and rewards to show that every payoff profile in W can be decomposed into payoffs from some action profile to be played currently and credible promises of future continuation payoffs which are crucially also in W. I then show that, with T-public (mediated) strategies, W is locally self-decomposable at some $(\underline{r}, \underline{\Delta})$ and remains so for all $(r, \Delta t) \leq (\underline{r}, \underline{\Delta})$.

8 Conclusion

This paper studies repeated games with frequent actions, secret monitoring and infrequent coordination, showing how to sustain dynamic equilibria with imperfect monitoring that converges to Brownian motion. The approach developed above relies on the use of mediated strategies, a plausible generalization of private strategies, which simplify the delay and dissemination of the arrival of endogenous strategic information. These mediated strategies may be thought of as dynamic information management institutions. These institutions form latent variables for each player that are revealed at regular intervals. The incentive schemes in this paper rely on empirical likelihood tests of obedience that not only apply to discrete-time problems, but also to continuous-time problems. The results and techniques seem general enough to apply generally, including environments with persistent as well as private payoff-relevant information. Continuous-time games can be useful for analyzing strategic outcomes with fixed discount rates. It would be interesting in the future to use them to understand the best forms of dynamic incentives for such fixed discount rates.

References

Abreu, D., P. Milgrom, and D. G. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713–33.

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory*, 39, 251–269.

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

Aoyagi, M. (2005): "Collusion Through Mediated Communication in Repeated Games with Imperfect Private Monitoring," *Economic Theory*, 25, 455–475.

Bernheim, B. D. and E. Madsen (2013): "Business Stealing in Imperfect Cartels," Working paper.

Bhaskar, V. (2000): "The robustness of repeated game equilibria to incomplete payoff information," University of Essex Working Paper.

Commission, E. (2001): "Commission Decision of 7 June 2000 (Case COMP/36.545/F3 — Amino Acids," *Official Journal of the European Communities*, 24–72.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597–626.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73, 377–415.

Faingold, E. and Y. Sannikov (2011): "Reputation in Continuous-Time Games," *Econometrica*, 79, 773–876.

Forges, F. (1986): "An approach to communication equilibria," *Econometrica*, 54, 1375–1385.

——— (1990): "Universal mechanisms," *Econometrica*, 1341–1364.

Fudenberg, D. and D. Levine (1994): "Efficiency and observability with long-run and short-run players," *Journal of Economic Theory*, 62, 103–135.

——— (2007): "Continuous time limits of repeated games with imperfect public monitoring," *Review of Economic Dynamics*, 10, 173–192.

——— (2009): "Repeated Games with Frequent Signals," *The Quarterly Journal of Economics*, 124, 233–265.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

Gneiting, T. and A. Raftery (2007): "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378.

Green, E. and R. Porter (1984): "Noncooperative collusion under imperfect price information," *Econometrica: Journal of the Econometric Society*, 87–100.

Harrington, Jr., J. E. and A. Skrzypacz (2011): "Private monitoring and communication in cartels: Explaining recent collusive practices," *American Economic Review*, 101, 2425–49.

Hexner, E. (1943): *The International Steel Cartel*, Chapel Hill.

Hoeffding, W. (1956): "On the distribution of the number of successes in independent trials," *The Annals of Mathematical Statistics*, 27, 713–721.

Hörner, J. and W. Olszewski (2006): "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, 74, 1499–1544.

Kandori, M. (2003): "Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring," *Econometrica*, 71, 345–353.

Kandori, M. and H. Matsushima (1998): "Private Observation, Communication, and Collusion," *Econometrica*, 66, 627–652.

Kandori, M. and I. Obara (2006): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," *Econometrica*, 74, 499–519.

Marshall, R. C. and L. M. Marx (2012): *The Economics of Collusion: Cartels and Bidding Rings*, MIT Press.

Myerson, R. (1986): "Multistage games with communication," *Econometrica*, 54, 323–358.

Obara, I. (2009): "Folk theorem with communication," *Journal of Economic Theory*, 144, 120–134.

Rahman, D. (2012a): "But Who Will Monitor the Monitor?" *American Economic Review*, 102, 2267–2297.

——— (2012b): "Mediating Collusion with Flexible Production," Mimeo.

——— (2013): "Information Delay in Games with Frequent Actions," Working paper.

Rahman, D. and I. Obara (2010): "Mediated Partnerships," *Econometrica*, 78, 285–308.

Samuels, S. (1965): "On the number of successes in independent trials," *The Annals of Mathematical Statistics*, 36, 1272–1278.

Sannikov, Y. (2007): "Games with imperfectly observable actions in continuous time," *Econometrica*, 75, 1285–1329.

Sannikov, Y. and A. Skrzypacz (2007): "Impossibility of collusion under imperfect monitoring with flexible production," *The American Economic Review*, 97, 1794–1823.

——— (2010): "The role of information in repeated games with frequent actions," *Econometrica*, 78, 847–882.

Sugaya, T. (2010): "Folk Theorem in Repeated Games with Private Monitoring," Tech. rep., mimeo.

Tomala, T. (2009): "Perfect Communication Equilibria in Repeated Games with Imperfect Monitoring," *Games and Economic Behavior*, 67, 682–694.

A Proofs

Proof of Lemma 1. A correlated strategy μ is *enforceable* with respect to $W \subset \mathbb{R}^2$ and δ if there exists a vector v and a function $w : A \times \Omega \to W$ such that

$$v_i = (1-\delta)u_i(\mu) + \delta \sum_{(a,\omega)} \Pr(\omega|a) w_i(a,\omega)\mu(a) \quad \forall i, \text{ and}$$

$$(1-\delta)\sum_{a_{-i}} \Delta u_i(a,b_i)\mu(a) + \delta \sum_{(a_{-i},\omega)} \Delta \Pr(\omega|a,b_i) w_i(a,\omega)\mu(a) \leq 0 \quad \forall (i,a_i,b_i).$$

The first family of equations above describes value recursion. The second family describes incentive compatibility: discouraging recommendation-contingent deviations.

Following Fudenberg et al. (1994), if μ is enforceable with respect to W and δ with the pair (v,w), we will say that w enforces μ with respect to v and δ, and that v is *decomposable* with respect to μ, W and δ. If μ is enforceable with respect to some W and δ, call it simply *enforceable*. Let $B(W,r,\Delta t)$ be the set of all decomposable payoff vectors as we vary μ with respect to fixed W, r and Δt. W is *self-decomposable* if $W \subset B(W,r,\Delta t)$ for some r.

A smooth[18] subset $W \subset \underline{U}$ is *decomposable on tangent hyperplanes* if for every point v on the boundary of W there exists a correlated strategy μ with finite support such that (i) $u(\mu)$ is separated from W by the (unique) hyperplane P_v that is tangent to W at u, and (ii) there exists a continuation payoff function $w : A \times \Omega \to P_v$ that enforces μ.

I will now argue decomposability on tangent hyperplanes, for the following reason.

Lemma 13 (Fudenberg et al, 1994, Theorem 4.1)**.** *If a smooth set $W \subset U_+$ is decomposable on tangent hyperplanes then $\underline{r} > 0$ exists with $W \subset E(r,\Delta t)$ for all $r < \underline{r}$, where, $E(r,\Delta t)$ is the set of public communication equilibrium payoffs.*

Let W be a smooth subset of the interior of the feasible, individually rational set $\underline{U}$. Let λ be the outward unit normal vector to W at v. For decomposability on tangent hyperplanes, I must show that given λ there is a correlated strategy μ such that (i) $\sum_i \lambda_i v_i < \sum_i \lambda_i u_i(\mu)$, and (ii) there exist continuation values w that enforce μ with $\sum_i \lambda_i w_i(a,\omega) = 0$ for all (a,ω). To decompose v, let $w_i^+(a)$ be the payment to player i after an up jump if the mediator recommended a, and similarly write $w_i^-(a)$ after a down jump. For simplicity, I assume that $w_i^+(a) + w_i^-(a) = 0$ for all a, and write $w_i(a) = 2w_i^+(a) = -2w_i^-(a)$. Let $w_i(D_i,D_{-i}) = 0$ and write $w_i(C_i,C_{-i}) = w_{i2}$, $w_i(C_i,D_{-i}) = -w_i(D_i,C_{-i}) = w_{i1}$.

If $\lambda \leq 0$, this is easy: v is decomposable into the pure strategy profile (D_1,D_2) and $w \equiv 0$.

If $\lambda_i = 0$ and $\lambda_j > 0$, decompose v into the pure strategy profile (C_i,D_j) and continuation values w as follows. First, player j needs no incentives to defect, so let $w_j \equiv 0$. On the other hand, enforceability for player i requires that $1-\delta \leq \delta w_{i1}(p_1 - p_0)$. Since $1-\delta \leq r\Delta t$ and $p_1 - p_0 = \frac{1}{2}(x_1 - x_0)\sqrt{\Delta t}$, it follows that $\delta w_{i1} \geq 2r\sqrt{\Delta t}/(x_1 - x_0)$ implies enforceability.

[18]Smooth means (i) closed and convex, (ii) with nonempty interior, and (iii) with a boundary that is a C^2-submanifold (Fudenberg et al, 1994, Definition 4.3).

For the remaining cases of λ, choose μ as follows:

	C	D
C	μ_0	μ_2
D	μ_1	

Let $\mu_0 + \mu_1 + \mu_2 = 1$ and every entry in the table above be strictly positive. Since W is closed and in the interior of $\underline{U}$, there exists $\varepsilon > 0$ such that for every vector $\lambda \not\leq 0$ there exists $\mu = (\mu_0, \mu_1, \mu_2) \geq \varepsilon$ such that $\sum_i \lambda_i u_i(\mu) > \sum_i \lambda_i v_i$. Indeed, if $\lambda_i > 0 \geq \lambda_j$ choose $\mu_0 = \mu_j = \varepsilon$ and $\mu_i = 1 - 2\varepsilon$, whereas if $\lambda \gg 0$ choose $\mu_0 = 1 - 2\varepsilon$ and $\mu_1 = \mu_2 = \varepsilon$.

With this notation, recommending cooperation is incentive compatible if

$$(1-\delta)(\mu_0 + \mu_j) \leq \delta[\mu_0(p_2 - p_1)w_{i2} + \mu_j(p_1 - p_0)w_{i1}].$$

Similarly, recommending defection requires the inequality

$$-(1-\delta) \leq \delta(p_2 - p_1)w_{i1}.$$

If $\lambda_i > 0 > \lambda_j$ and $x_2 > x_1$ then let $w_{i1} = w_{j1} = 0$, $\delta w_2 = 2r\sqrt{\Delta t}/[\varepsilon(x_2 - x_1)]$ and $\delta w_{i2} = \delta w_2 / |\lambda_i|$. Since $|\lambda_i| < 1$, incentive compatibility follows. Moreover, by construction $\lambda_i w_{i2} + \lambda_j w_{j2} = 0$, too. If $x_2 = x_1$, let $w_{i2} = 0$ for all i, choose $\delta w_1 = 2r\sqrt{\Delta t}/[\varepsilon(x_1 - x_0)]$ and let $\delta w_{i1} = \delta w_1/\lambda_i$. Incentive compatibility and budget balance again follow. Notice that these continuation values also yield incentive compatibility and budget balance if $\lambda \gg 0$. Therefore, W is decomposable on tangent hyperplanes, as claimed. It remains to show that for $\underline{\Delta} > 0$ sufficiently small, if $W \subset E(r, \underline{\Delta})$ then $W \subset E(r, \Delta t)$ for $\Delta t < \underline{\Delta}$. Following Fudenberg et al. (1994, proof of Theorem 4.1, p. 1035), choose any $v \in W$. Let μ and w decompose v as above. Change the coordinate system so that v is the origin, the first axis is the line connecting v to $u(\mu)$ and the remaining axes lie in P_v. For any vector x, write $x = (x^0, x^1)$, where x^0 is the component on the first axis and x^1 is the component in P_v. Since W is smooth, by Taylor's Theorem there exists $\delta^* < 1$, a constant $\hat{C} > 0$ and a neighborhood $\mathcal{O}$ of the origin such that, for all $\delta > \delta^*$, if $x \in \mathcal{O}$ then $\|x^1\| < \hat{C}\sqrt{(1-\delta)/\delta}$ and $x^0 \leq -\|u(\mu)\|(1-\delta)/\delta$ imply that x belongs to the interior of W. By decomposability on tangent hyperplanes, there exists (i) $r > 0$ such that $e^{-r\underline{\Delta}} > \delta^*$, and (ii) $\underline{w} \in \mathcal{O}$ (after the coordinate change) that enforces μ with $\|\underline{w}\| < \frac{1}{2}\hat{C}\sqrt{(1-\delta)/\delta}$. Think of the vector $\underline{w}$, belonging to P_v, as x^1 above, and $\underline{x}^0 = -\|u(\mu)\|(1-\delta)/\delta$, therefore $(\underline{x}^0, \underline{w})$ belongs to the interior of W. Now consider any $\Delta t < \underline{\Delta}$. Define $w = \underline{w}\sqrt{\Delta t/\underline{\Delta}}$ to be continuation values for $(r, \Delta t)$, and let $x^0 = -\|u(\mu)\|(1-\delta)/\delta$ with $\delta = e^{-r\Delta t}$. The scaled vector (x^0, w) enforces μ at $(r, \Delta t)$ and remains in the interior of W if $r > 0$ is small. That w enforces μ follows from the derivation of decomposability on tangent hyperplanes above. That (x^0, w) belongs to the interior of W for small $\underline{\Delta}$ follows because

$$w = \underline{w}\sqrt{\frac{\Delta t}{\underline{\Delta}}} < \tfrac{1}{2}\hat{C}\sqrt{\frac{r\Delta t}{r\underline{\Delta}}}\sqrt{\frac{1-e^{-r\underline{\Delta}}}{e^{-r\underline{\Delta}}}} \leq \tfrac{1}{2}\hat{C}\sqrt{\frac{1}{r\underline{\Delta}}}\sqrt{\frac{1-e^{-r\underline{\Delta}}}{e^{-r\underline{\Delta}}}}\sqrt{\frac{1-e^{-r\Delta t}}{e^{-r\Delta t}}} < \hat{C}\sqrt{\frac{1-e^{-r\Delta t}}{e^{-r\Delta t}}},$$

since $(1-e^{-r\underline{\Delta}})/e^{-r\underline{\Delta}} < 4r\underline{\Delta}$ for small enough $r > 0$. Therefore, w belongs to the interior of W for all $\Delta t < \underline{\Delta}$. This completes the proof of Lemma 1. □

Proof of Lemma 3. CI fails if and only if for some (i, a_i, ω) there is a vector $y_i \geq 0$ such that $\Pr(\omega|a) = \sum_{b_i \neq a_i} y_i(b_i) \Pr(\omega|b_i, a_{-i})$ for all a_{-i}. Let $\sigma_i' = y_i / \sum_{b_i} y_i(b_i)$, and for any $\pi \in (0,1)$, let $\sigma_i(a_i) = 1 - \pi$ and $\sigma_i(b_i) = \pi\sigma_i'(b_i)$ for $b_i \neq a_i$. Write $\Pr(\omega|\sigma_i, a_{-i}) = \sum_{b_i} \sigma_i(b_i) \Pr(\omega|b_i, a_{-i})$. Now, as was claimed, CI fails if and only if there exist (i, a_i, ω) and σ_i such that

$$\frac{\Pr(\omega|\sigma_i, a_{-i})}{\Pr(\omega|a)} = \frac{\Pr(\omega|\sigma_i, b_{-i})}{\Pr(\omega|a_i, b_{-i})} \qquad \forall (a_{-i}, b_{-i}).$$

With private monitoring, UI fails if and only if for some (i, a_i, ω_i) there exists $y_i \geq 0$ such that $\Pr(\omega|a) = \sum_{(a_i', \omega_i') \neq (a_i, \omega_i)} y_i(a_i', \omega_i') \Pr(\omega_i', \omega_{-i}|a_i', a_{-i})$ for all (a_{-i}, ω_{-i}). Now define a deviation just as before, with $\hat{\sigma}_i' = y_i / \sum_{(a_i', \omega_i')} y_i(a_i', \omega_i')$. The rest of the proof is the same as that for CI. □

Proof of Lemma 4. Suppose that for every $\underline{\Delta} > 0$ there exists $\Delta t \in (0, \underline{\Delta})$ such that conic independence fails: either $p(a_i, \cdot) \in \text{cone}\{p(b_i, \cdot) : b_i \neq a_i\}$ or $q(a_i, \cdot) \in \text{cone}\{q(b_i, \cdot) : b_i \neq a_i\}$ for some (i, a_i). Without loss, the second inclusion holds infinitely often at (i, a_i) as $\underline{\Delta} \to 0$, so there exists $y_{i\Delta t}(a_i, \cdot) \geq 0$ such that for every a_{-i},

$$\begin{aligned} q(a) &= \sum_{b_i \neq a_i} y_{i\Delta t}(a_i, b_i) p(b_i, a_{-i}) \quad \Leftrightarrow & (15) \\ \tfrac{1}{2}[1 - x(a)\sqrt{\Delta t}] &= \tfrac{1}{2} \sum_{b_i \neq a_i} y_{i\Delta t}(a_i, b_i)[1 - x(b_i, a_{-i})\sqrt{\Delta t}] \quad \Leftrightarrow & (16) \\ x(a) &= \sum_{b_i \neq a_i} y_{i\Delta t}(a_i, b_i) x(b_i, a_{-i}) + (1 - \sum_{b_i \neq a_i} y_{i\Delta t}(a_i, b_i))/\sqrt{\Delta t}. & (17) \end{aligned}$$

For any sequence $\{\Delta t_m > 0\}$ decreasing to 0, consider the corresponding sequence $\{y_{im} = y_{i\Delta t_m}\}$. Let $\overline{y}_{im}(a_i) = \sum_{b_i \neq a_i} y_{im}(a_i, b_i)$. Without loss assume that $\{\overline{y}_{im}(a_i)\}$ is a monotone sequence, so it has a (possibly infinite) limit, $\overline{y}_i(a_i)$. If $\overline{y}_i(a_i) = 0$ then obviously (16) fails. If $\overline{y}_i(a_i) = \infty$ then divide (17) by $\overline{y}_{im}(a_i)$, rearrange terms, and let $\pi_{im}(a_i) \in \Delta(A_i \setminus \{a_i\})$ be defined according to $\pi_{im}(a_i, b_i) = y_{im}(a_i, b_i)/\overline{y}_{im}(a_i)$ to obtain

$$0 = \sum_{b_i \neq a_i} \pi_{im}(a_i, b_i) x(b_i, a_{-i}) - \frac{x(a)}{\overline{y}_{im}(a_i)} + \frac{\overline{y}_{im}(a_i) - 1}{\overline{y}_{im}(a_i)} / \sqrt{\Delta t_m}. \tag{18}$$

But the right-hand side of (18) explodes, a contradiction. To see this, note that the first term, $\sum_{b_i \neq a_i} \pi_{im}(a_i, b_i) x(b_i, \cdot)$, lies in the bounded set $\text{conv}\{x(b_i, \cdot) : b_i \neq a_i\}$, so is bounded, too. The second term, $x(a_i, \cdot)/\overline{y}_{im}(a_i)$, clearly converges to zero because $\overline{y}_{im}(a_i) \to \infty$. The third term explodes, too, since $[\overline{y}_{im}(a_i) - 1]/\overline{y}_{im}(a_i) \to 1$ and $1/\sqrt{\Delta t_m} \to \infty$. If $\overline{y}_i(a_i) \in \mathbb{R}_+$ is different from 1 then again $[1 - \overline{y}_{im}(a_i)]/\sqrt{\Delta t_m}$ explodes, leading to another contradiction of (18). Finally, suppose that $\overline{y}_i(a_i) = 1$. Since $y \geq 0$, it follows that $y_{im}(a_i, b_i)$ is a bounded sequence, hence has a convergent subsequence. Taking subsequences of subsequences if necessary, there is a subsequence such that all $y_{im}(a_i, b_i)$ converge together to some limit $y_i(a_i, b_i)$. Depending on the rate at which $\overline{y}_{im}(a_i) \to 1$ relative to $\Delta t_m \to 0$, the term $(1 - \overline{y}_{im}(a_i))/\sqrt{\Delta t_m}$ can converge to any real number, independently of a_{-i}. The claim now follows from (17), since it implies that (4) fails. □

Proof of Proposition 1. First notice that, by definition of x and p,

$$\begin{aligned}\frac{\Pr(\omega|a)}{\Pr(\omega|\sigma_i,a_{-i})}-\frac{\Pr(\omega|a_i,b_{-i})}{\Pr(\omega|\sigma_i,b_{-i})} &= 1+\frac{\Pr(\omega|a)}{\Pr(\omega|\sigma_i,a_{-i})}-1-\frac{\Pr(\omega|a_i,b_{-i})}{\Pr(\omega|\sigma_i,b_{-i})}\\ &= \frac{\Pr(\omega|a)-\Pr(\omega|\sigma_i,a_{-i})}{\Pr(\omega|\sigma_i,a_{-i})}-\frac{\Pr(\omega|a_i,b_{-i})-\Pr(\omega|\sigma_i,b_{-i})}{\Pr(\omega|\sigma_i,b_{-i})}.\end{aligned}$$

If x fails conditional identifiability then there is a mixed strategy $\hat{\sigma}_i$ and a scalar α_i with

$$x(a)-\sum_{b_i\neq a_i}\hat{\sigma}_i(b_i)x(b_i,a_{-i})=\alpha_i \qquad \forall a_{-i}.$$

Applying the previous formula to $\hat{\sigma}_i$ yields

$$\frac{\Pr(\omega|a)-\Pr(\omega|\hat{\sigma}_i,a_{-i})}{\Pr(\omega|\hat{\sigma}_i,a_{-i})}-\frac{\Pr(\omega|a_i,b_{-i})-\Pr(\omega|\hat{\sigma}_i,b_{-i})}{\Pr(\omega|\hat{\sigma}_i,b_{-i})}=\pm\left[\frac{\frac{1}{2}\alpha_i\sqrt{\Delta t}}{\Pr(\omega|\hat{\sigma}_i,a_{-i})}-\frac{\frac{1}{2}\alpha_i\sqrt{\Delta t}}{\Pr(\omega|\hat{\sigma}_i,b_{-i})}\right].$$

Therefore,

$$\frac{1}{\sqrt{\Delta t}}\left[\frac{\Pr(\omega|a)}{\Pr(\omega|\hat{\sigma}_i,a_{-i})}-\frac{\Pr(\omega|a_i,b_{-i})}{\Pr(\omega|\hat{\sigma}_i,b_{-i})}\right]=\pm\frac{\frac{1}{2}\alpha_i[\Pr(\omega|\hat{\sigma}_i,b_{-i})-\Pr(\omega|\hat{\sigma}_i,a_{-i})]}{\Pr(\omega|\hat{\sigma}_i,a_{-i})\Pr(\omega|\hat{\sigma}_i,b_{-i})}\to 0 \text{ as } \Delta t\to 0.$$

Conversely, if x satisfies conditional identifiability then for every mixed strategy σ_i such that $\sigma_i(a_i)<1$ there exist (a_{-i},α_i) and (b_{-i},β_i) such that $\alpha_i\neq\beta_i$ and both

$$x(a)-\sum_{b_i}\sigma_i(b_i)x(b_i,a_{-i})=\alpha_i \quad \text{and} \quad x(a_i,b_{-i})-\sum_{b_i}\sigma_i(b_i)x(b)=\beta_i.$$

Following the previous steps,

$$\begin{aligned}\frac{1}{\sqrt{\Delta t}}\left[\frac{\Pr(\omega|a)}{\Pr(\omega|\hat{\sigma}_i,a_{-i})}-\frac{\Pr(\omega|a_i,b_{-i})}{\Pr(\omega|\hat{\sigma}_i,b_{-i})}\right] &= \pm\frac{\frac{1}{2}\alpha_i\Pr(\omega|\hat{\sigma}_i,b_{-i})-\beta_i\Pr(\omega|\hat{\sigma}_i,a_{-i})}{\Pr(\omega|\hat{\sigma}_i,a_{-i})\Pr(\omega|\hat{\sigma}_i,b_{-i})}\\ &\to \pm(\alpha_i-\beta_i) \quad \text{as} \quad \Delta t\to 0,\end{aligned}$$

since each of the probabilities above converges to $\frac{1}{2}$. □

Proof of Lemma 5. Every deviation $\tilde{\sigma}_i$ from $\tilde{\mu}$ corresponds to a deviation from μ that does not depend on public announcements, so, integrating out α, it follows that

$$\begin{aligned}U_i(\mu|\tilde{\sigma}_i) &= (1-\delta)\sum_{(\tau,a^\tau,b_i^\tau,\omega^\tau,\alpha^\tau)}\delta^{\tau-1}u_i(b_{i\tau},a_{-i\tau})\Pr(a^\tau,\omega^\tau,\alpha^\tau,b_i^\tau|\mu,\tilde{\sigma}_i)\\ &= (1-\delta)\sum_{(\tau,a^\tau,b_i^\tau,\omega^\tau)}\delta^{\tau-1}u_i(b_{i\tau},a_{-i\tau})\Pr(a^\tau,\omega^\tau,b_i^\tau|\tilde{\mu},\tilde{\sigma}_i) \;=\; U_i(\tilde{\mu}|\tilde{\sigma}_i)\geq U_i(\tilde{\mu}).\end{aligned}$$

Therefore, $\tilde{\mu}$ is a private equilibrium. □

Proof of Proposition 2. Let μ be a public mediated strategy that discourages public deviations. Let σ_i be any deviation from μ, not necessarily public. The probability of $(a^\tau, \omega^\tau, b_i^\tau)$ equals

$$\Pr(a^\tau, \omega^\tau, b_i^\tau | \mu, \sigma_i) = \prod_{\rho=1}^{\tau} \mu(a_\rho | a^{\rho-1}, \omega^{\rho-1}) \Pr(\omega_\rho | b_{i\rho}, a_{-i\rho}) \sigma_i(b_{i\rho} | a_{i\rho}, b_i^{\rho-1}, \omega^{\rho-1}, a^{\rho-1})$$

as usual. Since μ is public, for $s \leq \tau$, the probability conditional on (a^s, ω^s, b_i^s) equals

$$\Pr(a^\tau, \omega^\tau, b_i^\tau | a^s, \omega^s, b_i^s, \mu, \sigma_i) = \Pr(a^\tau, \omega^\tau, b_i^\tau | \mu, \sigma_i) / \Pr(a^s, \omega^s, b_i^s | \mu, \sigma_i).$$

Let $s+1$ be the first period where σ_i depends on b_i^s. Decompose $U_i(\mu|\sigma_i)$ as follows:

$$U_i(\mu|\sigma_i) = (1-\delta^s) U_{is}(\mu|\sigma_i) + \delta^s \sum_{(a^s, \omega^s, b_i^s)} U_i^s(\mu | a^s, \omega^s, b_i^s, \sigma_i) \Pr(a^s, \omega^s, b_i^s | \mu, \sigma_i),$$

where $U_{is}(\mu|\sigma_i) = \frac{1-\delta}{1-\delta^s} \sum_{\tau=1}^{s} \sum_{(a^\tau, \omega^\tau, b_i^\tau)} \delta^\tau u_i(a_\tau) \Pr(a^\tau, \omega^\tau, b_i^\tau | \mu, \sigma_i)$ and

$$U_i^s(\mu | h_i^{s+1}, \sigma_i) = (1-\delta) \sum_{h_i^{\tau+1} > h_i^{s+1}} \delta^{\tau-s-1} u_i(a_\tau) \Pr(h_i^\tau | h_i^{s+1}, \mu, \sigma_i(h_i^{s+1})).$$

Let $b_i^{s*}(h_0^{s+1}, b_i^s) \in \arg\max_{\hat{b}_i^s} U_i^s(\mu | h_0^{s+1}, b_i^s, \sigma_i(h_0^{s+1}, \hat{b}_i^s))$. It is easy to see that $b_i^{s*}(h_0^{s+1}, b_i^s) = b_i^{s*}(h_0^{s+1})$ does not depend on b_i^s. Indeed, letting $b_{is+1}^\tau = (b_{is+1}, \ldots, b_{i\tau})$,

$$\begin{aligned}
U_i^s(\mu | h_0^{s+1}, b_i^s, \sigma_i(h_0^{s+1}, \hat{b}_i^s)) &= (1-\delta) \sum_{h_i^{\tau+1} > h_i^{s+1}} \delta^{\tau-s-1} u_i(a_\tau) \Pr(h_i^\tau | h_0^{s+1}, b_i^s, \mu, \sigma_i(h_0^{s+1}, \hat{b}_i^s)) \\
&= (1-\delta) \sum_{h_i^{\tau+1} > h_i^{s+1}} \delta^{\tau-s-1} u_i(a_\tau) \times \\
&\quad \prod_{\rho=s+1}^{\tau} \mu(a_\rho | a^{\rho-1}, \omega^{\rho-1}) \Pr(\omega_\rho | b_{i\rho}, a_{-i\rho}) \sigma_i(b_{i\rho} | a_{i\rho}, b_{is+1}^{\rho-1}, \hat{b}_i^s, \omega^{\rho-1}, a^{\rho-1}) \\
&= U_i^s(\mu | h_0^{s+1}, \sigma_i(h_0^{s+1}, \hat{b}_i^s)).
\end{aligned}$$

Therefore, a deviation is optimal regardless of the private part of a player's history. Letting $\sigma_i^s = \sigma_i$ for $\tau \leq s$ and $\sigma_i^s(h_i^{\tau+1}) = \sigma_i(h_0^{\tau+1}, b_{is+1}^t, b_i^{s*}(h_0^{s+1}))$, it follows that $\sigma_i^s(h_i^{s+1}) = \sigma_i^s(h_0^{s+1})$ is a public deviation up to and including period $s+1$, and moreover $U_i^s(\mu | h_i^{s+1}, \sigma_i) \leq U_i^s(\mu | h_i^{s+1}, \sigma_i^s)$. Now let $s'+1$ be the next period where σ_i^s is not public, and repeat the algorithm above to obtain $\sigma_i^{s'}$. Proceeding inductively, the limiting deviation is public and its value exceeds that of σ_i. $\square$

Proof of Proposition 3. I will prove the claim for conditional irreversibility both in probabilities and drifts. Assuming λ is regular, consider the following linear program.

$$
\begin{aligned}
V_\lambda(\mu,\Delta t) \;&=\; \sup_{\gamma,\xi\geq 0,\beta,\pi} \gamma \;\text{ s.t. }\; \xi_i(a,\omega)\mu(a)\leq\mu(a) \quad \forall(i,a,\omega),\\
\gamma\sum_{a_{-i}}\Delta u_i(b_i,a_{-i})\mu(a) \leq \sum_{(a_{-i},\omega)}&[\tfrac{\lambda_i}{|\lambda_i|}\xi_i(a,\omega)+\beta_i(a,\omega)]\Delta\Pr(\omega|a,b_i)\mu(a) \quad \forall(i,a_i,b_i),\\
0 &\leq \sum_{a_{-i}}(\xi_i(a,\omega)-\pi)\Pr(\omega|a_{-i},b_i)\mu(a) \quad \forall(i,a_i,b_i,\omega),\\
\pi &\geq \sum_{(i,a,\omega)}\xi_i(a,\omega)\Pr(\omega|a)\mu(a),\\
&\sum_{i=1}^n\lambda_i\beta_i(a,\omega)=0 \quad \forall(a,\omega).
\end{aligned}
$$

If a proper λ-balanced scoring rule exists then $V_\lambda(\mu,\Delta t)>0$. The dual of this problem is

$$
\begin{aligned}
V(\mu,\Delta t) = \sup_{\eta,\sigma,y\geq 0}\sum_{(i,a,\omega)}\eta_i(a,\omega)\mu(a) \;&\text{ s.t. }\; \sum_{i=1}^n\Delta u_i(\mu,\sigma_i)\geq 1,\\
\eta_i(a,\omega)\geq\tfrac{\lambda_i}{|\lambda_i|}\Delta\Pr(\omega|a,\sigma_i)+\sum_{b_i}y_i(a_i,b_i,\omega)&\Pr(\omega|a_{-i},b_i)-\hat{y}\Pr(\omega|a) \quad \forall(i,a,\omega),\\
\hat{y} = \sum_{(i,a,b_i,\omega)}&y_i(a_i,b_i,\omega)\Pr(\omega|a_{-i},b_i)\mu(a),\\
\Delta\Pr(\omega|a,\sigma_i) &= \lambda_i\hat{\eta}(a,\omega) \quad \forall(a,\omega).
\end{aligned}
$$

With the same argument as in Proposition 6, $V_\lambda=0$ implies that

$$
\tfrac{\lambda_i}{|\lambda_i|}\Delta\Pr(\omega|a,\sigma_i)+\sum_{b_i}y_i(a_i,b_i,\omega)\Pr(\omega|a_{-i},b_i)-\hat{y}\Pr(\omega|a)=0 \quad \forall(i,a,\omega).
$$

The first dual constraint requires that the deviation profile σ_i be profitable. The last one requires that σ_i be λ-unattributable. Finally. the equation above requires that σ be conditionally reversible with respect to λ. The proof for conditional irreversibility in drifts is similar; just replace $\frac{\lambda_i}{|\lambda_i|}\Delta\Pr(\omega|a,\sigma_i)$ in the dual with $\frac{\lambda_i}{|\lambda_i|}\Delta x(a,\sigma_i)$. □

Proof of Proposition 4. The proof that UI is generic follows from Rahman (2012a, Theorem 8). The proof for CI is almost identical. □

Proof of Lemma 8. By belief stability, $\sum_{a_{-i}}\xi_i(a,\omega)\Pr(\omega|a)\mu(a)\geq\pi_i\sum_{a_{-i}}\Pr(\omega|a)\mu(a)$ for all (i,a_i,ω). Summing with respect to (a_i,ω) yields $\sum_{(a,\omega)}\xi_i(a,\omega)\Pr(\omega|a)\mu(a)\geq\pi_i$. Belief stability also requires that $\pi_i=\sum_{(a,\omega)}\xi_i(a,\omega)\Pr(\omega|a)\mu(a)$, which is violated if there exists (i,a_i,ω) such that $\sum_{a_{-i}}\xi_i(a,\omega)\Pr(\omega|a)\mu(a)>\pi_i\sum_{a_{-i}}\Pr(\omega|a)\mu(a)$.

With private monitoring, belief stability (5) together with the argument above yields

$$
\sum_{(a_{-i},\omega_{-i})}\xi_i(a,\omega)\Pr(\omega|a)\mu(a)\geq\pi_i\sum_{(a_{-i},\omega_{-i})}\Pr(\omega|a)\mu(a)
$$

instead. This establishes (i). For (ii), let ξ be any proper scoring rule with prior punishment probability vector π. If $\pi_i > \alpha_i$, scale down ξ_i to obtain $\xi_i'(a,\omega) = \alpha_i \xi_i(a,\omega)/\pi_i$. It follows that

$$\pi_i' = \sum_{(a,\omega)} \xi_i'(a,\omega) \Pr(\omega|a)\mu(a) = \alpha_i \sum_{(a,\omega)} \xi_i(a,\omega) \Pr(\omega|a)\mu(a)\pi_i = \alpha_i \pi_i/\pi_i = \alpha_i.$$

Finally, if $\pi_i < \alpha_i$ then pick $\beta_i = (\alpha_i - \pi_i)/(1-\pi_i)$ and $\xi_i'(a,\omega) = \beta_i + (1-\beta_i)\xi_i(a,\omega)$. Notice that $\beta_i \in (0,1)$, so $\xi_i' \in [0,1]$ and still satisfies (10). Now, similarly to the previous case where $\pi_i > \alpha_i$, we obtain that $\pi_i' = \beta_i + (1-\beta_i)\pi_i = \alpha_i$. Since the transformations applied to ξ were affine and monotone, it follows that ξ' is still a proper scoring rule. The same proof applies with private monitoring, except that incentive constraints are those in (6) instead of (10). □

Proof of Proposition 6. A simple application of duality. Consider the following linear program:

$$\begin{aligned} V(\mu,\Delta t) &= \sup_{\gamma,\xi\geq 0,\pi} \gamma \text{ s.t. } \xi_i(a,\omega)\mu(a) \leq \mu(a) \quad \forall (i,a,\omega), \\ \gamma \sum_{a_{-i}} \Delta u_i(b_i,a_{-i})\mu(a) &\leq \sum_{(a_{-i},\omega)} \xi_i(a,\omega) \Delta \Pr(\omega|a,b_i)\mu(a) \quad \forall (i,a_i,b_i), \\ \sum_{a_{-i}} \Pr(\omega|a_{-i},b_i)\mu(a)\pi_i &\leq \sum_{a_{-i}} \xi_i(a,\omega) \Pr(\omega|a_{-i},b_i)\mu(a) \quad \forall (i,a_i,b_i,\omega), \\ \pi_i &\geq \sum_{(a,\omega)} \xi_i(a,\omega) \Pr(\omega|a)\mu(a) \quad \forall i. \end{aligned}$$

If a proper scoring rule ξ exists then $V(\mu,\Delta t) > 0$. The dual of this problem is given by

$$\begin{aligned} V(\mu,\Delta t) &= \inf_{\eta,\sigma,y\geq 0} \sum_{(i,a,\omega)} \eta_i(a,\omega)\mu(a) \text{ s.t. } \sum_i \Delta u_i(\mu,\sigma_i) \geq 1, \\ \eta_i(a,\omega) &\geq \Delta \Pr(\omega|a,\sigma_i) + \sum_{b_i} y_i(a_i,b_i,\omega) \Pr(\omega|a_{-i},b_i) - \hat{y}_i \Pr(\omega|a) \quad \forall (i,a,\omega), \\ \hat{y}_i &= \sum_{(a_i,b_i,\omega)} y_i(a_i,b_i,\omega) \Pr(\omega|a_{-i},b_i)\mu(a) \quad \forall i. \end{aligned}$$

Suppose $V(\mu,\Delta t) = 0$. Since $\eta \geq 0$, necessarily $\eta_i(a,\omega) = 0$ for all (i,a,ω). Substituting for $\hat{y}_i$, it follows that $\sum_{(a,\omega)} \sum_{b_i} y_i(a_i,b_i,\omega) \Pr(\omega|a_{-i},b_i)\mu(a) - \hat{y}_i \Pr(\omega|a)\mu(a) = 0$. Since, in addition, $\sum_\omega \Delta \Pr(\omega|a,\sigma_i) = 0$ for each a, the right-hand side of the second dual inequality above adds up to zero with respect to (a,ω). So, if there exists (a,ω) such that this right-hand side is negative then there must exist another (a,ω) for which this right-hand side is positive, contradicting the hypothesis that $\eta_i(a,\omega) = 0$ for all (a,ω). Therefore, this right-hand side must equal zero for all (a,ω). Rearranging this equation and dividing by $\mu(a) > 0$ yields

$$\sum_{b_i} \Pr(\omega|a_{-i},b_i)[\sigma_i(b_i|a_i) + y_i(a_i,b_i,\omega)] - \Pr(\omega|a)[\textstyle\sum_{b_i}\sigma_i(b_i|a_i) + \hat{y}_i] = 0 \quad \forall (i,a,\omega). \tag{19}$$

Finally, dividing by $\sum_{b_i}\sigma_i(b_i|a_i) + \hat{y}_i$ we obtain that $\Pr(\omega|a)$ is a positive linear combination of $\Pr(\omega|a_{-i},b_i)$, with weights that depend on (a_i,b_i,ω), but not on a_{-i}. By the first dual inequality above, σ_i is (proportional to) a profitable deviation, so Pr fails conditional identifiability.

With private monitoring, simply replace the second and third families of inequalities in the primal problem above with (5) and (6). The rest of the argument follows almost identically, except for a few minor changes. First, the dual problem becomes

$$V(\mu,\Delta t) = \inf_{\eta,\sigma,y\geq 0} \sum_{(i,a,\omega)} \eta_i(a,\omega)\mu(a) \quad \text{s.t.} \quad \sum_i \Delta u_i(\mu,\sigma_i) \geq 1,$$

$$\eta_i(a,\omega) \geq \Delta \Pr(\omega|a,\sigma_i) + \sum_{(a_i',\omega_i')} y_i(a_i,\omega_i,a_i',\omega_i') \Pr(\omega_i',\omega_{-i}|a_{-i},a_i') - \hat{y}_i \Pr(\omega|a) \quad \forall (i,a,\omega),$$

$$\hat{y}_i = \sum_{(a_i,\omega_i,a_i',\omega_i')} y_i(a_i,\omega_i,a_i',\omega_i') \Pr(\omega_i',\omega_{-i}|a_{-i},a_i')\mu(a) \quad \forall i,$$

where $\Delta \Pr(\omega|a,\sigma_i) = \Pr(\omega|a,\sigma_i) - \Pr(\omega|a)$,

$$\Pr(\omega|a,\sigma_i) = \sum_{(a_i',\omega_i')} \Pr(\omega_i',\omega_{-i}|a_i',a_{-i})\sigma_i(a_i',\omega_i|a_i,\omega_i') \quad \text{and} \quad \sigma_i(a_i',\omega_i|a_i,\omega_i') = \sum_{\{\rho_i:\rho_i(\omega_i')=\omega_i\}} \sigma_i(a_i',\rho_i|a_i).$$

Apart from replacing the left-hand side of (19) above with

$$\sum_{(a_i',\omega_i')} \Pr(\omega_i',\omega_{-i}|a_{-i},a_i')[\sigma_i(a_i',\omega_i|a_i,\omega_i') + y_i(a_i,\omega_i,a_i',\omega_i')] - \Pr(\omega|a)[\textstyle\sum_{(a_i',\omega_i')}\sigma_i(a_i',\omega_i|a_i,\omega_i') + \hat{y}_i]$$

the rest of the proof follows identically. □

Proof of Lemma 9. Follows immediately from the Maximum Theorem and the fact that $C_i^* < \infty$ while ξ is a proper scoring rule. The fact that μ is a completely mixed correlated strategy is used only to obtain continuity for π_i^{**}, since it is defined with conditional probabilities given by μ. □

Proof of Proposition 7. Let $W(\mu,\Delta t) = V(\mu,\Delta t)/\sqrt{\Delta t}$ for all $\Delta t > 0$, where $V(\mu,\Delta t)$ was defined in the proof of Proposition 6, so

$$W(\mu,\Delta t) \;=\; \sup_{\lambda,\xi,\pi} \lambda \quad \text{s.t.} \quad 0 \leq \xi_i(a,\omega) \leq 1 \quad \forall (i,a,\omega),$$

$$\lambda \sum_{a_{-i}} \Delta u_i(b_i,a_{-i})\mu(a) \leq \tfrac{1}{2} \sum_{(a_{-i},\omega)} \xi_i(a,\omega)\omega\Delta x(a,b_i)\mu(a) \quad \forall (i,a_i,b_i),$$

$$\sum_{a_{-i}} \Pr(\omega|a_{-i},b_i)\mu(a)\pi_i \leq \sum_{a_{-i}} \xi_i(a,\omega)\Pr(\omega|a_{-i},b_i)\mu(a) \quad \forall (i,a_i,b_i,\omega),$$

$$\pi_i = \sum_{(a,\omega)} \xi_i(a,\omega)\Pr(\omega|a)\mu(a) \quad \forall i.$$

By Lemma 4, $W(\mu,\Delta t) > 0$ for all small $\Delta t > 0$. (Note that W may still be unbounded.) I will show that $W(\mu,0) = \lim_{\Delta t\to 0} W(\mu,\Delta t) > 0$. This clearly yields $\overline{C}_i^* > 0$ and $\overline{\pi}_i \in (0,1)$, since $\overline{\pi}_i = 0$ or 1 implies $\overline{\xi}_i(a,\omega)$ all equal 0 or 1 by virtue of the fact that $\overline{\pi}_i = \frac{1}{2}\sum_{(a,\omega)} \xi_i(a,\omega)\mu(a)$. By the incentive constraint above, this in turn implies $\lambda = 0$, contradicting $W(\mu,0) > 0$.

The dual of the problem above equals

$$W(\mu,\Delta t) = \inf_{\eta,\sigma,y\geq 0} \sum_{(i,a,\omega)} \eta_i(a,\omega)\mu(a) \quad \text{s.t.} \quad \sum_i \Delta u_i(\mu,\sigma_i) \geq 1,$$

$$\eta_i(a,\omega) \geq \omega \Delta x(a,\sigma_i) + \sum_{b_i} y_i(a_i,b_i,\omega)\Pr(\omega|a_{-i},b_i) - \hat{y}_i \Pr(\omega|a) \quad \forall (i,a,\omega),$$

$$\hat{y}_i = \sum_{(a,b_i,\omega)} y_i(a_i,b_i,\omega)\Pr(\omega|a_{-i},b_i)\mu(a) \quad \forall i.$$

For a contradiction, suppose that $W(\mu,\Delta t) \to 0$, and assume that (η,σ,y) solve the dual at $\Delta t > 0$. By the Maximum Theorem, $W(\mu,0) = 0$. Taking a subsequence if necessary, as $\Delta t \to 0$ the solution (η,σ,y) converges to $(\overline{\eta},\overline{\sigma},\overline{y})$, say, which must satisfy the dual constraints evaluated at $\Delta t = 0$ and $\overline{\eta} = 0$, since $\overline{\eta} \geq 0$:

$$\sum_i \Delta u_i(\mu,\overline{\sigma}_i) \geq 1,$$

$$0 \geq \pm \Delta x_i(a,\overline{\sigma}_i) + \tfrac{1}{2}\left[\sum_{b_i} \overline{y}_{i\pm}(a_i,b_i) - \overline{\hat{y}}_i\right] \quad \forall (i,a), \tag{20}$$

$$\overline{\hat{y}}_i = \tfrac{1}{2}\sum_{(a,b_i)} \mu(a)[\overline{y}_{i+}(a_i,b_i) + \overline{y}_{i-}(a_i,b_i)] \quad \forall i. \tag{21}$$

Substituting for $\overline{\hat{y}}_i$ from (21) and adding all the inequalities in (20) weighted by μ yields

$$0 \geq \sum_{a\in A} \mu(a)[\Delta x_i(a,\overline{\sigma}_i) - \Delta x_i(a,\overline{\sigma}_i)] + \tfrac{1}{2}\sum_{a\in A}\mu(a)[\sum_{b_i} \overline{y}_{i+}(a_i,b_i) + \overline{y}_{i-}(a_i,b_i) - \overline{\hat{y}}_i] = 0.$$

Hence, every right-hand side in (20) equals zero. This clearly contradicts CI-x.

For the general case with public monitoring, replace $\omega\Delta x(a,\sigma_i)$ above with $\Delta \Pr(\omega|a,\sigma_i)/\sqrt{\Delta t}$. If $W(\mu,\Delta t) \to 0$ then $\Delta \Pr(\omega|a,\sigma_i)/\sqrt{\Delta t} + \sum_{b_i} y_i(a_i,b_i,\omega)\Pr(\omega|a_{-i},b_i) - \hat{y}_i \Pr(\omega|a) \to 0$ for all (i,a,ω). Replace y_i with $\tilde{y}_i = y_i/\sqrt{\Delta t}$ to get $\Pr(\omega|a,(\sigma_i + \tilde{y}_i)/\sqrt{\Delta t}) - ((\hat{\sigma}_i + \tilde{\hat{y}}_i)/\sqrt{\Delta t})\Pr(\omega|a) \to 0$, where $\hat{\sigma}_i = \sum_{b_i} \sigma_i(b_i|a_i)$ and, without loss, this right-hand side does not depend on a_i. Let $\tilde{\sigma}_i(a_i,b_i,\omega) = \sigma_i(b_i|a_i) + y_i(a_i,b_i,\omega)$. Dividing by $\sum_{b_i} \tilde{\sigma}_i(a_i,b_i,\omega)\Pr(\omega|a)$ and subtracting the same condition for b_{-i} from that for a_{-i} yields $[\mathrm{Lr}(\omega|a,\tilde{\sigma}_i) - \mathrm{Lr}(\omega|a_i,b_{-i},\tilde{\sigma}_i)]/\sqrt{\Delta t} \to 0$. As $\Delta t \to 0$, $\tilde{\sigma}_i$ remains proportional to a deviation for some i by the first dual constraint even in the limit (with factor of proportionality bounded away from 0), which contradicts Assumption 3. The proof for the general case with private monitoring is almost identical, therefore omitted. □

Proof of Lemma 10. We already established that $(1-\delta)\Delta u_i(\sigma_i) \leq e^{-rc} w_i \Delta\pi_i(\sigma_i) f_{iT-1}(\tau_i^{**})$. Since $1-\delta \leq r\Delta t$, it follows that $r\Delta t \Delta u_i(\sigma_i) \leq e^{-rc} w_i \Delta\pi_i(\sigma_i) f_{iT-1}(\tau_i^{**})$ implies (11). Since $T \leq c/\Delta t$, this is implied by

$$rce^{rc}\frac{\Delta u_i(\sigma_i)}{\Delta \pi_i(\sigma_i)} \leq w_i T f_{iT-1}(\tau_i^{**})$$

after rearrangement. Maximizing the left-hand side with respect to σ_i and substituting for f_i,

$$rce^{rc}\frac{\Delta u_i^*}{\Delta \pi_i^*} \leq w_i T \binom{T-1}{\tau_i^{**}} \pi_i^{T-1-\tau_i^{**}} (1-\pi_i)^{\tau_i^{**}}.$$

But $\frac{T}{\pi_i}\binom{T-1}{\tau_i^{**}} = \frac{T-\tau_i^{**}}{\pi_i}\binom{T}{\tau_i^{**}} \geq \frac{1+\pi_i^{**}(T-1)}{\pi_i}\binom{T}{\tau_i^{**}} \geq T\binom{T}{\tau_i^{**}}$, from which the claim follows. □

Proof of Lemma 11. The probability mass functions $f_{iT-1}(\cdot|h_i^\tau)$ and $f_{iT-1}(\cdot)$ are obtained from the convolution of independent Bernoulli trials. By belief stability (9), each of the Bernoulli trials that generate $f_{iT-1}(\cdot|h_i^\tau)$ have a failure probability that is greater than or equal to the corresponding Bernoulli trial that generates $f_{iT-1}(\cdot)$, with failure rate π_i. At the same time, these failure probabilities cannot be larger than π_i^{**} by definition of π_i^{**}.

Let us simplify notation for the purpose of the proof. I will use f instead of the more cumbersome f_{iT-1}. I will use the following classic observation due to Samuels (1965) regarding the mode of a sequence of independent Bernoulli trials. (The notation is slightly different because π_i corresponds to the probability of failure—not success.)

Lemma 14 (Samuels, 1965, Theorem 1)**.** *If k is an integer such that*

$$k \leq (1-\overline{\pi}_i)(T-1) \leq k+1,$$

where $\overline{\pi}_i = \sum_{\tau=1}^{T-1} \pi_{i\tau}/(T-1)$ and $\pi_i^\vee = \max_\tau\{\pi_{i\tau}\}$, then $f^\vee(k) \geq f^\vee(k-1)$, where $f^\vee$ is the probability mass function that excludes a trial with failure probability $\pi_i^\vee$.

Thus, the mode of a Binomial with failure probability p and n trials is $\lfloor (n+1)p \rfloor$. Pick any sequence of failure probabilities in order from highest to lowest: $\{\pi_{i\tau}^1\}$. Let f be the probability mass function of successes from independent Bernoulli trials with failure probabilities $\{\pi_{i\tau}^1\}$. Let f_1 be defined by

$$f_1(n) = f^\vee(n-1) + \pi_i \Delta f^\vee(n),$$

where $\Delta f^\vee(n) = f^\vee(n) - f^\vee(n-1)$. Of course, $f(n)$ has the same expression except for $\pi_i^{1\vee}$ replacing π_i. If k_1 satisfies the conditions of Lemma 14 above for $\{\pi_{it}^1\}$ then $f_1(k_1) \leq f(k_1)$, since $\pi_i \leq \pi_i^{1\vee}$ and $\Delta f^\vee(k_1) \geq 0$.

Now let $\{\pi_{i\tau}^2\}$ be the sequence of failure probabilities obtained by first replacing π_{i1}^1 with π_i and then reordering the probabilities from highest to lowest. Let f_2 be defined by

$$f_2(n) = f_1^\vee(n-1) + \pi_i \Delta f_1^\vee(n).$$

Similarly, if k_2 satisfies the conditions of Lemma 14 above for $\{\pi_{i\tau}^2\}$ then $f_2(k_2) \leq f_1(k_2)$. But since $\pi_{i\tau}^2 \leq \pi_{i\tau}^1$ for all τ, it follows that $k_2 \leq k_1$, therefore $f_2(k_1) \leq f_1(k_1) \leq f(k_1)$. Proceeding inductively, it follows that $\pi_{i\tau}^{T-1} = \pi_i$ for all τ and $f_{T-1}(k_1) \leq f(k_1)$. In other words, the probability of k_1 successes with failure probability π_i at every period t is less than or equal to that with arbitrary failure probabilities $\{\pi_{i\tau}^1\}$.

Finally, the largest possible value of k_1 is associated with $\pi_{i\tau}^1 = \pi_i^{**}$ for all τ, rendering $k_1 = \tau_i^{**}$. Notice that, by unimodality, for any other sequence of $\pi_{i\tau}^1$'s between π_i and π_i^{**}, $k_1 \leq \tau_i^{**}$, therefore $f^\vee(k_1) \geq f^\vee(k_1-1)$ implies that $f^\vee(\tau_i^{**}) \geq f^\vee(\tau_i^{**}-1)$. This finally implies the claim. □

Proof of Proposition 8. The first claim follows by a simple application of the de Moivre-Laplace Theorem. Recall the condition from Lemma 10:

$$rce^{rc}\frac{\Delta u_i^*}{\Delta\pi_i^*} \leq w_i T\binom{T}{\tau_i^{**}}\pi_i^{T-\tau_i^{**}}(1-\pi_i)^{\tau_i^{**}}.$$

Rearranging $\sqrt{T} \leq \sqrt{c/\Delta t}$ and writing $\Delta\pi_i^* = z_i^* \frac{1}{2}\sqrt{\Delta t}$ yields

$$rce^{rc}\frac{\Delta u_i^*}{z_i^*\sqrt{c}} \leq w_i \tfrac{1}{2}\sqrt{T}\binom{T}{\tau_i^{**}}\pi_i^{T-\tau_i^{**}}(1-\pi_i)^{\tau_i^{**}}.$$

Since $\tau_i^{**} = \lfloor(1-\pi_i^{**})(T-1)\rfloor$ and by assumption $\pi_i = \frac{1}{2}$ along the sequence of scoring rules as $\Delta t \to 0$, by the de Moivre-Laplace Theorem,

$$\begin{aligned}
\tfrac{1}{2}\sqrt{T}\binom{T}{\tau_i^{**}}\pi_i^{T-\tau_i^{**}}(1-\pi_i)^{\tau_i^{**}} &\approx \frac{\frac{1}{2}}{\sqrt{2\pi\pi_i(1-\pi_i)}}e^{-\frac{(\tau_i^{**}-(1-\pi_i)T)^2}{2T\pi_i(1-\pi_i)}} \\
&\approx \frac{1}{\sqrt{2\pi}}e^{-\frac{[(\pi_i-\pi_i^{**})\sqrt{T}]^2}{2\pi_i(1-\pi_i)}} \\
&\approx \frac{1}{\sqrt{2\pi}}e^{-\frac{(\overline{z}_i^{**}\frac{1}{2}\sqrt{c})^2}{2\pi_i(1-\pi_i)}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{(\overline{z}_i^{**}\sqrt{c})^2}{2}} = \varphi(\overline{z}_i^{**}\sqrt{c}).
\end{aligned}$$

Again by the de Moivre-Laplace Theorem, the left-hand side converges to the right-hand side, meaning that the ratio of the left and right hand sides above converges to 1.

For the second claim, choose w_i so that inequality (13) holds with equality. Substituting this expression into the lifetime utility for player i yields

$$v_i \to u_i(\mu) - \frac{rc\Delta\overline{u}_i^*}{1-e^{-rc}}\frac{\Pi_{i0}}{\varphi(\overline{z}_i^{**}\sqrt{c})\overline{z}_i^*\sqrt{c}} \quad \text{as } \Delta t \to 0.$$

By the Central Limit Theorem and a similar calculation to the one above, $\Pi_{i0} \to 1 - \Phi(\overline{z}_i^{**}\sqrt{c})$. Moreover, recent improvements to the Berry-Esseen Theorem,[19] establish uniform convergence in the Central Limit Theorem of order $\frac{1}{2}\sqrt{\Delta t/c}$. For the third claim, the hazard rate of the normal distribution explodes faster than linearly, so its inverse implodes:

$$\frac{1-\Phi(\overline{z}_i^{**}\sqrt{c})}{\varphi(\overline{z}_i^{**}\sqrt{c})\overline{z}_i^*\sqrt{c}} \to 0 \quad \text{as } c \to \infty.$$

Finally, for the last claim, if $c \to \infty$ but $rc \to 0$ as $r \to 0$ (e.g., $c = r^{\varepsilon-1}$ and $0 < \varepsilon < 1$) then $rc/(1-e^{-rc}) \to 1$, therefore

$$\lim_{r\to 0}\lim_{\Delta t\to 0} v_i = u_i(\mu).$$

In the general case beyond Example 3, it is possible that $z_i^{**} \to \infty$. If not, then the previous argument survives intact. If so, then—instead of fixing c—fix $T \in \mathbb{N}$ arbitrarily large and let

$$w_i \geq \frac{r\sqrt{T\Delta t}}{\varphi(2\Delta\pi_i^*\sqrt{T})}\frac{\Delta u_i^*}{z_i^*}.$$

By Proposition 7, $z_i^* \not\to 0$ as $\Delta t \to 0$, so w_i is well defined and bounded above. The rest of the arguments now follow similarly, except with the different interpretation of T fixed instead of c. □

[19]See http://en.wikipedia.org/wiki/Berry-Esseen_theorem.

Proof of Lemma 12. The proof is straightforward. The first derivative of $\hat{R}_i$ equals

$$\begin{aligned}
\hat{R}_i'(\hat{\pi}_i) &= \sum_{\tau=0}^{\tau_{i*}} \binom{T}{\tau}[\tau\hat{\pi}_i^{\tau-1}(1-\hat{\pi}_i)^{T-\tau} - (T-\tau)\hat{\pi}_i^{\tau}(1-\hat{\pi}_i)^{T-1-\tau}] \\
&= T\sum_{\tau=0}^{\tau_{i*}} \binom{T-1}{\tau-1}\hat{\pi}_i^{\tau-1}(1-\hat{\pi}_i)^{T-\tau} - \binom{T-1}{\tau}\hat{\pi}_i^{\tau}(1-\hat{\pi}_i)^{T-1-\tau} \\
&= -T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}}(1-\hat{\pi}_i)^{T-1-\tau_{i*}} < 0,
\end{aligned}$$

therefore $\hat{R}_i'(\hat{\pi}_i)$ is a decreasing function of $\hat{\pi}_i$. Similarly, the second derivative equals

$$\begin{aligned}
\hat{R}_i''(\hat{\pi}_i) &= -T\binom{T-1}{\tau_{i*}}[\tau_{i*}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-1-\tau_{i*}} - (T-1-\tau_{i*})\hat{\pi}_i^{\tau_{i*}}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}] \\
&= -T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}[\tau_{i*}(1-\hat{\pi}_i) - (T-1-\tau_{i*})\hat{\pi}_i] \\
&= -T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}[\tau_{i*} - (T-1)\hat{\pi}_i] \\
&= -T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}[\lfloor \pi_i(T-1)\rfloor - 1 - (T-1)\hat{\pi}_i] \\
&\geq -T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}[\pi_i(T-1) - 1 - (T-1)\hat{\pi}_i] \\
&= T\binom{T-1}{\tau_{i*}}\hat{\pi}_i^{\tau_{i*}-1}(1-\hat{\pi}_i)^{T-2-\tau_{i*}}[(T-1)(\hat{\pi}_i - \pi_i) + 1] > 0.
\end{aligned}$$

Therefore, $\hat{R}_i$ is strictly convex on $[\pi_i, 1]$. □

Proof of Proposition 10. The first two claims follow immediately from Proposition 9 above, since $R_{i0} = \frac{1}{2}$. It remains to establish the third claim, which, since $rc/(1-e^{-rc}) \to 1$ as $rc \to 0$, boils down to showing that $D_i < \infty$ and $\frac{1}{2}D_i \to \Delta u_i^d$.

To see that $D_i < \infty$, writing $E[\Delta\hat{u}_i(\sigma_i^T)]$, defined just after Lemma 12 as just $E[\Delta\hat{u}_i]$, by convexity

$$D_{iT}(\sigma_i^T) := \frac{E[\Delta\hat{u}_i]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]} = \frac{E[\Delta\hat{u}_i]}{E[\Delta\hat{\pi}_i]}\frac{E[\Delta\hat{\pi}_i]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]} \leq \frac{E[\Delta\hat{u}_i]}{E[\Delta\hat{\pi}_i]}E\left[\frac{\Delta\hat{\pi}_i}{R_i(\pi_i) - R_i(\hat{\pi}_i)}\right].$$

Notice that $\frac{E[\Delta\hat{u}_i]}{E[\Delta\hat{\pi}_i]} \leq \frac{\Delta u_i^*}{\Delta\pi_i^*} < \infty$. By Lemma 12, the slope $\frac{R_i(\pi_i)-R_i(\hat{\pi}_i)}{\hat{\pi}_i-\pi_i}$ is a decreasing function of $\hat{\pi}_i$, with minimum at π_i^{**}. Therefore, $\frac{\Delta\hat{\pi}_i}{R_i(\pi_i)-R_i(\hat{\pi}_i)} \leq \frac{\Delta\pi_i^{**}}{R_i(\pi_i)-R_i(\pi_i^{**})} < \infty$. Moreover, by Lemma 9 and Proposition 7, $\overline{C}_i^* = \lim_{\Delta t\to 0} C_i^* < \infty$, and

$$\frac{z_i^{**}}{R_i(\pi_i) - R_i(\pi_i^{**})} \to \frac{\overline{z}_i^{**}}{\Phi(0) - \Phi(-\overline{z}_i^{**}\sqrt{c})} \quad \text{as } \Delta t \to 0$$

by the central limit theorem, so

$$\lim_{\Delta t\to 0} D_{iT}(\sigma_i^T) \leq \frac{\Delta\overline{u}_i^*}{\overline{z}_i^*}\frac{\overline{z}_i^{**}}{\Phi(0) - \Phi(-\overline{z}_i^{**}\sqrt{c})} < \infty$$

unless $\overline{z}_i^{**} = 0$, which would imply that μ is implementable with completely constant transfers, that is, there is no incentive problem. This last inequality implies that $D_i < \infty$.

For the last claim, notice first that $D_i \geq \Delta u_i^d/[\Phi(0) - \Phi(-z_i^d\sqrt{c})]$ by having player i always deviate to σ_i^d. It follows that $\lim_{c\to\infty} D_i \geq 2\Delta u_i^d$. For the converse inequality, let

$$\tilde{\sigma}_i^T \in \arg\max_{\sigma_i^T} \frac{E[\Delta u_i(\sigma_i^T)]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]}$$

be an optimal dynamic deviation for a T-period block. There exists $\pi_i^\vee \in [\pi_i, \pi_i^{**}]$ such that

$$E[R_i(\hat{\pi}_i)] = R_i(\pi_i^\vee),$$

therefore $E[R_i(\pi_i) - R_i(\hat{\pi}_i)] = R_i(\pi_i) - R_i(\pi_i^\vee)$. By convexity of R_i, $\pi_i^\vee \leq E[\hat{\pi}_i]$, but since R_i is strictly decreasing, $\pi_i^\vee = \pi_i$ only if $E[\Delta u_i(\sigma_i^T)] \leq 0$. Let $\tilde{\pi}_i$ be the average failure probability generated by $\tilde{\sigma}_i$, that is, $\tilde{\pi}_i = \frac{1}{T}E[\sum_\tau \pi_{i\tau}]$, where $\pi_{i\tau}$ is the (random) failure probability associated with $\tilde{\sigma}_i^T$ in period τ. Let $\tilde{z}_i$ be the failure drift associated with $\tilde{\pi}_i$. Notice that $E[\Delta\hat{u}_i] \leq \Delta\hat{u}_i^*(\tilde{z}_i)$, where $\Delta u_i^+(\sigma_i) = \sum_{(a_i,b_i)} \sigma_i(b_i|a_i)\max\{\sum_{a_{-i}} \mu(a)\Delta u_i(a, b_i), 0\}$ and

$$\Delta\hat{u}_i^*(\tilde{z}_i) := \max_{\sigma_i \geq 0} \{\Delta u_i^+(\sigma_i) : \tilde{z}_i \geq \sum_{(a,\omega)} \xi_i(a,\omega)\omega\Delta x(a,b_i)\sigma_i(b_i|a_i)\mu(a),\ \sum_{b_i} \sigma_i(b_i|a_i) \leq 1\ \forall a_i\}.$$

This follows simply because $E[\Delta\hat{u}_i]$ is constructed as a deterministic average of T random utility payoffs (one for each time period) that can be incorporated into σ_i. The function $\Delta\hat{u}_i^*(\tilde{z}_i)$ is defined as the value of a linear program. Its dual is given by

$$\min_{\lambda,\kappa_i\geq 0}\{\lambda z_i + \sum_{a_i} \kappa_i(a_i) : \Delta u_i^+(a_i,b_i) \leq \lambda \sum_{a_{-i}} \xi_i(a,\omega)\omega\Delta x(a,b_i)\mu(a) + \kappa_i(a_i)\},$$

where $\Delta u_i^+(a_i, b_i) = \max\{\sum_{a_{-i}} \mu(a)\Delta u_i(a, b_i), 0\}$. To estimate D_i, let us first bound the directional derivative of $\Delta\hat{u}_i^*$ at 0. By duality, this is bounded above by λ^*, where (λ^*, κ_i^*) solve the dual above. Since the dual minimizes its objective, this λ^* is bounded above by

$$C_i^+ = \sup_{\sigma_i \in M(A_i)} \frac{\sum_{(a_i,b_i)} \Delta u_i^+(a_i,b_i)\sigma_i(b_i|a_i)}{\sum_{a,b_i} \xi_i(a,\omega)\omega\Delta x(a,b_i)\sigma_i(b_i|a_i)\mu(a)}.$$

This follows because C_i^+ is clearly a feasible solution for λ in the dual together with $\kappa_i \equiv 0$. By Proposition 7, there exists a proper scoring rule for Δu_i^+, so without loss we may assume that ξ_i is such a rule. Therefore, $C_i^+ = C_i^* < \infty$ and $\overline{C}_i^* < \infty$.

As $\Delta t \to 0$, $R_i(\pi_i) - R_i(\pi_i^\vee) \to \Phi(0) - \Phi(-\overline{z}_i^\vee\sqrt{c})$, where $\overline{z}_i^\vee \in [0, z_i^{**}]$ is the limit of failure drifts corresponding to $\pi_i^\vee$ as $\Delta t \to 0$. If $\overline{z}_i^\vee = 0$ then by l'Hopital's Rule

$$\lim_{\Delta t\to 0} \max_{\sigma_i^T} \frac{E[\Delta u_i(\sigma_i^T)]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]} \leq \frac{\overline{C}_i^*}{\varphi(0)\sqrt{c}} = \frac{\Delta u_i^{+*}}{\varphi(0)\overline{z}^*\sqrt{c}}.$$

By convexity of $\Phi(z)$ for $z \leq 0$ (see Figure 5), it follows that

$$\varphi(0) \geq \frac{\Phi(0) - \Phi(-\overline{z}_i^*\sqrt{c})}{\overline{z}_i^*\sqrt{c}}.$$

Therefore,

$$\frac{\Delta u_i^{+*}}{\varphi(0)\overline{z}^*\sqrt{c}} \leq \frac{\Delta u_i^{+*}}{\Phi(0) - \Phi(-\overline{z}_i^*\sqrt{c})}.$$

Taking the limit as $c \to \infty$ yields $2\Delta u_i^{+*}$, where $\Delta u_i^{+*} \leq \Delta u_i^d$ by definition of Δu_i^d.

Taking a subsequence if necessary, $\overline{z}_i^\vee$ converges to some failure drift $\overline{\overline{z}}_i^\vee \in [0, z_i^{**}]$ as $c \to \infty$. If this limit is different from zero then $E[\Delta u_i(\overline{\overline{z}}_i^\vee)] \geq 0$ and $\Phi(0) - \Phi(-\overline{z}_i^\vee \sqrt{c}) \to \frac{1}{2}$, so $\lim_{c\to\infty} D_i \leq 2\Delta u_i^d$. On the other hand, now suppose instead that $\overline{\overline{z}}_i^\vee = 0$. If $\overline{z}_i^\vee \to 0$ as fast as $1/\sqrt{c}$ or slower then trivially $\lim_{c\to\infty} D_i = 0$, as the numerator tends to zero but not the denominator (recall that $E[\Delta \hat{u}_i] \leq \Delta \hat{u}_i^*$ defined above), which contradicts the previously derived lower bound on D_i. If convergence is faster, again by l'Hopital's rule

$$\frac{E[\Delta \hat{u}_i]}{E[R_i(\pi_i) - R_i(\hat{\pi}_i)]} \approx \frac{\Delta u_i^{+*}}{\varphi(0)\overline{z}^*\sqrt{c}} \to 0.$$

This again contradicts our previously derived lower bound. The result now follows. In the general case beyond Example 3, the proof is amended just as for the proof of Proposition 8. □

Proof of Theorem 2. The proof is a simple application of Fudenberg et al. (1994). First I study the setting of Example 3, then I discuss the general cases of public and private monitoring.

To begin, let W be any smooth subset of the interior of U_0.[20] I will eventually show that W is locally self-decomposable. Let $B(W, r, \Delta t, c)$ be the set of all payoff vectors $v \in U$ such that, for $T = \lfloor c/\Delta t \rfloor$, there is a T-public mediated strategy $\tilde{\mu}$ and continuation payoffs $w : (A \times \Omega)^T \to W$ that *enforce* v, i.e., (i) v_i equals the expected lifetime utility for player i from playing $\tilde{\mu}$ for the first T-period block followed by the expected continuation payoff $E[w_i]$, and (ii) the mechanism $(\tilde{\mu}, w)$ is incentive compatible in the sense that every $\sigma_i^T \in \mathcal{M}_i^T$ is unprofitable. The set $B(W, r, \Delta t, c)$ of so-called *decomposable* payoffs is indexed by the set of feasible continuation payoffs W, the discount rate r, the time-step Δt and the length of calendar time c of the T-period blocks during which equilibrium strategies are private. This leads to the following version of local self-decomposability.

Definition 14. Given $\Delta t > 0$ and $c > 0$, $W \subset \mathbb{R}^n$ is called *self-decomposable* if $W \subset B(W, r, \Delta t, c)$ for some $r > 0$. W is *locally* self-decomposable if for each $v \in W$ there exists $r > 0$ and an open set $\mathcal{O}$ containing v such that $\mathcal{O} \cap W \subset B(W, r, \Delta t, c)$.

This standard definition of local self-decomposability leads to the following useful lemma, which is proved in Lemma 4.2 of Fudenberg et al. (1994).

Lemma 15. *Fix $\Delta t > 0$ and $c > 0$. If $W \subset \mathbb{R}^n$ is compact, convex and locally self-decomposable then there exists $\underline{r} > 0$ such that $W \subset E(r, \Delta t, c)$ for all $r \in (0, \underline{r}]$, where $E(r, \Delta t, c)$ is the set of T-public communication equilibrium payoffs when the time interval has length Δt, $T = \lfloor c/\Delta t \rfloor$ and players' common discount rate is $r > 0$.*

[20] By definition (Fudenberg et al, 1994, p. 1012), the set W is *smooth* if (i) it is closed and convex, (ii) it has nonempty interior, and (iii) its boundary is a C^2-submanifold of $\mathbb{R}^n$.

The key step towards proving Theorem 2 is the following.

Lemma 16. *Every smooth subset of the interior of V^{**} is locally self-decomposable.*

Proof. Let W be any such smooth subset of the interior of V^{**} and pick any $v \in W$. If $v \in \text{int } W$ (as in Fudenberg et al, 1994), choose an open ball $\mathcal{O}$ containing v whose closure belongs to the interior of W. Then there exists $\delta < 1$ such that for each $u \in \mathcal{O}$ there exists $u' \in W$ such that $u = (1-\delta^T)u(\mu_0) + \delta^T u'$, where μ_0 is a correlated equilibrium of the stage game. Now let $r_{\mathcal{O}}(\overline{c})$ solve $\delta = e^{-r\overline{c}}$, where $\overline{c}$ is determined later.

Next, suppose that $v \in \partial W$.[21] Since W is in the interior of V^{**}, there exists $\varepsilon_1, \varepsilon_2 > 0$ such that W belongs to the interior of $V_*^{\varepsilon_1,\varepsilon_2} \cap U_0$, too. Therefore, we may use our previously derived punishment and reward schemes consistently. Let $\Lambda = \{\lambda \in \mathbb{R}^n : \|\lambda\| = 1\}$ be the set of directions. Let $J = \{i \in I : \lambda_i \geq 0\}$, and choose $\mu \in \Delta^{\varepsilon_1}(A)$ such that $v_i \geq u_i^d(\mu) + \varepsilon_2$ for all $i \notin J$ and $u_i(\mu) \geq v_i + \varepsilon_2$ for all $i \in J$, therefore $u(\mu) \in \text{int } \underline{V}_J^{\varepsilon_1,\varepsilon_2}$.

Players in J face punishment schemes and the rest reward schemes. The mediator makes recommendations according to the correlated strategy μ independently during every period of a T-period block, where $T = \lfloor c/\Delta t \rfloor$ will be chosen later. Given μ, the scoring rule is determined in two steps. First, find a proper scoring rule ξ' that solves the linear program defined in the proof of Proposition 7 with value $W(\mu, \Delta t)$. By Assumption 3, $W(\mu, \Delta t) > 0$. Next, apply an affine transformation to ξ' to obtain a proper scoring rule ξ such that $\pi_i = \frac{1}{2}$ (Lemma 8(ii)). Since $\Delta^{\varepsilon_1}(A)$ is compact and $W(\mu, \Delta t)$ is continuous on that set, every $\mu \in \Delta^{\varepsilon_1}(A)$ has a well-defined proper scoring rule for each $\Delta t > 0$, and these converge to their corresponding limiting proper scoring rules as $\Delta t \to 0$. In each period, a scoring trial is implemented, and players' scores are calculated according to the rules of Section 7.

Consider two kinds of direction λ in trying to decompose v with respect to some μ. First, say $\lambda_i \neq 0$ for every i. If $\lambda_i > 0$, by Proposition 8 and its proof $\delta^T/(1-\delta^T)\Pi_{i0}w_i = u_i(\mu) - v_i$ and for every $\varepsilon_3 > 0$ there exists $\underline{\Delta} > 0$ such that

$$\frac{\delta^T}{1-\delta^T}\Pi_{i0}w_i > \frac{\Delta\overline{u}_i^*[1 - \Phi(\overline{z}_i^{**}\sqrt{c}) + \sqrt{\underline{\Delta}/c}]}{\overline{z}_i^*\sqrt{c}\varphi(\overline{z}_i^{**}\sqrt{c})(1-\varepsilon_3)}. \tag{22}$$

I will now derive a uniform bound on the right-hand side above. First of all, let

$$\overline{C}^* = \max_{i,\mu\in\Delta^{\varepsilon_1}(A)} \Delta\overline{u}_i^*/\overline{z}_i^*.$$

By the Maximum Theorem, $\overline{C}_i^*(\mu)$ is continuous on $\Delta^{\varepsilon_1}(A)$, a compact set, therefore its maximum is attained, so this maximization is well defined. Given $\varepsilon_3 > 0$, by continuity of $\overline{z}^*$ and $\overline{z}^{**}$ with respect to μ, the same choice of $\underline{\Delta}$ satisfies (22) in a neighborhood of μ. Repeating this exercise for every μ yields an open cover of $\Delta^{\varepsilon_1}(A)$ indexed by μ, with each neighborhood in the cover having its own associated $\underline{\Delta} > 0$. Since $\Delta^{\varepsilon_1}(A)$ is compact there exists a finite sub-cover. Let $\Delta' > 0$ be the minimum of the $\underline{\Delta}$'s in the finite sub-cover. Define the highest possible failure rate

[21] The ∂ notation stands for boundary, thus ∂W is the boundary of W.

by $z' = \max_{(i,\mu)}\{\overline{z}_i^{**} : \mu \in \Delta^{\varepsilon_1}(A)\}$, which is clearly finite. The tightest failure rate, for each $\underline{\Delta}$ and c, is defined by

$$z^* \in \arg\max_{z\in[0,z']} \frac{[1-\Phi(z\sqrt{c})+\sqrt{\underline{\Delta}/c}]}{\varphi(z\sqrt{c})\sqrt{c}}.$$

Substituting the value recursion expression above into the incentive constraint, and recognizing that $u_i(\mu) - v_i \geq \varepsilon_2$, yields a sufficient condition, uniform in i and μ, for incentive compatibility of punishment schemes:

$$\varepsilon_2 > \overline{C}^* \frac{[1-\Phi(z^*\sqrt{c})+\sqrt{\underline{\Delta}/c}]}{\varphi(z^*\sqrt{c})\sqrt{c}(1-\varepsilon_3)},$$

where $\varepsilon_3 \in (0,1)$ is arbitrary. As $\underline{\Delta} \to 0$ and $c \to \infty$, the right-hand side tends to zero, and as $c \to 0$ it tends to ∞. Therefore, there exists $c' > 0$ that satisfies this inequality, hence implies incentive compatibility for our punishment schemes based on any $\mu \in \Delta^{\varepsilon_1}(A)$.

If $\lambda_i < 0$ then, by Proposition 10, incentive compatibility is implied by

$$\frac{\delta^T}{1-\delta^T} R_{i0} w_i \geq \tfrac{1}{2} D_i,$$

where value recursion for rewards yields $\delta^T/(1-\delta^T)R_{i0}w_i = v_i - u_i(\mu)$. Recognizing that $v_i - u_i(\mu) \geq \Delta u_i^d + \varepsilon_2$ yields the sufficient condition

$$\varepsilon_2 > \tfrac{1}{2} D_i - \Delta u_i^d.$$

By the proof of Proposition 10, there exist $\hat{\Delta}$ small enough and $\hat{c}$ large enough to satisfy this inequality. Now fix $\underline{\Delta} = \min\{\Delta', \hat{\Delta}\}$ and $\overline{c} = \max\{c', \hat{c}\}$.

Recall $\delta^T/(1-\delta^T)\Pi_{i0}w_i = u_i(\mu) - v_i$ for i with $\lambda_i > 0$ and $\delta^T/(1-\delta^T)R_{i0}w_i = v_i - u_i(\mu)$ for i with $\lambda_i < 0$. Let $w_i = \alpha[u_i(\mu) - v_i]/\Pi_{i0}$ when $\lambda_i > 0$ and $w_i = \alpha[v_i - u_i(\mu)]/R_{i0}$ when $\lambda_i < 0$, for some $\alpha > 0$. Let $v_i' = v_i - w_i$ if $\lambda_i > 0$ and $v_i' = v_i + w_i$ if $\lambda_i < 0$. Now choose α such that $v' \in \operatorname{int} W$. Finally, choose $r_v > 0$ to solve $e^{-r\overline{c}}/(1-e^{-r\overline{c}}) = 1/\alpha$. It now follows that v is decomposable with respect to μ, W and $(r_v, \underline{\Delta}, \overline{c})$. Moreover, a small enough perturbation of v within W, with corresponding changes in w_i to preserve value recursion, maintains decomposability, $v' \in \operatorname{int} W$, and incentive compatibility (since the sufficient incentive inequalities above were strict) for the same r_v. Therefore, there is an open set $\mathcal{O}$ containing v such that $\mathcal{O} \cap W \subset B(W, r_v, \underline{\Delta}, \overline{c})$.

It remains to argue the case where $\lambda_i = 0$ for some i. For self-generation, we must amend other players' continuation value (see Figure 4 for geometric intuition). If $\lambda_i > 0$, he will now face a punishment scheme where his contingent payoffs are $v_i - \lambda_i\varepsilon_2(1-e^{-r\overline{c}})/e^{-r\overline{c}}$ if no punishment ensues an $v_i - \lambda_i\varepsilon_2(1-e^{-r\overline{c}})/e^{-r\overline{c}} - w_i$ if punishment ensues. Similarly, if $\lambda_i < 0$, he will now face a reward scheme where his contingent payoffs are $v_i + \lambda_i\varepsilon_2(1-e^{-r\overline{c}})/e^{-r\overline{c}}$ if no reward ensues an $v_i + \lambda_i\varepsilon_2(1-e^{-r\overline{c}})/e^{-r\overline{c}} + w_i$ if reward ensues. If $\lambda_i = 0$ then player i faces a punishment scheme, shifted by the amount $\Pi_{i0}w_i$. That is, if punishment ensues (which happens with probability Π_{i0}), player i's continuation payoff becomes $v_i + \Pi_{i0}w_i - w_i$, otherwise it becomes $v_i + \Pi_{i0}w_i$. Hence, player i's expected continuation payoff equals v_i. The incentive compatibility constraints derived above still hold relative to $J = \{i : \lambda_i \geq 0\}$.

Now, in order for the continuation values to be self-generated we choose w and $r > 0$ as follows. By value recursion, $e^{-r\overline{c}}/(1-e^{-r\overline{c}})\Pi_{i0}w_i = u_i(\mu) - v_i - \frac{1}{2}\lambda_i\varepsilon_2$ when $\lambda_i > 0$, $e^{-r\overline{c}}/(1-e^{-r\overline{c}})R_{i0}w_i = v_i - u_i(\mu) + \frac{1}{2}\lambda_i\varepsilon_2$ when $\lambda_i < 0$, and $v_i = u_i(\mu)$ for $\lambda_i = 0$. However, for $\lambda_i = 0$, the punishments and rewards for player i relative to v_i are less than or equal to w_i in magnitude. The incentive constraint for this player i with $\lambda_i = 0$ is given by

$$e^{-r\overline{c}}/(1-e^{-r\overline{c}})\Pi_{i0}w_i > \overline{C}^* \frac{[1-\Phi(z^*\sqrt{\overline{c}}) + \sqrt{\underline{\Delta}/\overline{c}}]}{\varphi(z^*\sqrt{\overline{c}})\sqrt{\overline{c}}(1-\varepsilon_3)} =: K.$$

Therefore, $e^{-r\overline{c}}/(1-e^{-r\overline{c}})w_i = K/\Pi_{i0}$. For $\lambda_i = 0$, choose $w_i = \alpha K/\Pi_{i0}$. For $\lambda_i > 0$, choose $w_i = \alpha[u_i(\mu) - v_i - \frac{1}{2}\lambda_i\varepsilon_2]/\Pi_{i0}$. For $\lambda_i < 0$, choose $w_i = \alpha[v_i - u_i(\mu) + \frac{1}{2}\lambda_i\varepsilon_2]/R_{i0}$.

It remains to show that for some small $\alpha > 0$, the continuation values v' belong to int W. For $\lambda_i \neq 0$ this is clear. For $\lambda_i = 0$, notice that the ratio of contingent payments to player i relative to transfers to everyone else (which, since λ is a unit vector, have length $\frac{1}{2}\varepsilon_2$) equals $\frac{K}{\Pi_{i0}\frac{1}{2}\varepsilon_2} < \infty$. Since W is smooth, a second-order Taylor series expansion of ∂W shows that, following the proof of Theorem 4.1 in Fudenberg et al. (1994), since

$$w_i = (K/\Pi_{i0})(1-e^{-r\overline{c}})/e^{-r\overline{c}} < \sqrt{\tfrac{1}{2}\varepsilon_2(1-e^{-r\overline{c}})/e^{-r\overline{c}}}$$

for sufficiently small $r > 0$, the change in continuation value induced by w_i is insufficient to escape from W. Therefore, it is possible to pick α, hence r, such that every $v' \in \text{int } W$. Let $r_v > 0$ be such an interest rate. Therefore, v is decomposable with respect to μ, W and $(r_v, \underline{\Delta}, \overline{c})$. Finally, as before, a small perturbation of v is still decomposable with the same parameters, so there is an open set $\mathcal{O}$ containing v such that $\mathcal{O} \cap W \subset B(W, r_v, \underline{\Delta}, \overline{c})$. □

For each $v \in W$, we just argued that there is an open set $\mathcal{O}_v$ and $(r, \Delta t, c)$ such that $\mathcal{O} \cap W \subset B(W, r, \Delta t, c)$. By compactness, the open cover $\{\mathcal{O}_v\}$ of W has a finite sub-cover, $\{\mathcal{O}_k : k \in \{1, \ldots, m\}\}$. Let $\underline{\Delta} = \min\{\Delta t_k\} > 0$, let $\overline{c} = \max\{c_k\} < \infty$ and finally define $\underline{r} = \min\{r_k c_k\}/\overline{c} < \min\{r_k\}$. By construction, $\underline{r}\overline{c} \leq r_k c_k$ for every k. Hence, $W \subset E(\underline{r}, \underline{\Delta}, \overline{c})$ for some $(\underline{r}, \underline{\Delta}, \overline{c})$. By construction of the open cover above and Propositions 8(1) and 10(1), $W \subset E(r, \Delta t, \overline{c})$ for all $(r, \Delta t) \leq (\underline{r}, \underline{\Delta})$. This proves Theorem 2 for the case of Example 3. In the general case beyond Example 3, the proof is amended just as for the proofs of Propositions 8 and 10. □

Proof of Theorem 3. Almost nothing changes from the outline of Example 6. Pick a profile v of payoffs that belongs to the interior of U_0. Let $\mu \in \Delta(A)$ be a completely mixed correlated strategy whose payoff profile strictly dominates v. Now the method of Example 6 can be applied by finding a profile of continuation value changes w on the line crossing $u(\mu)$ and v and letting $w = \alpha(v - u_0)$ be the expected change in payoffs from playing according to equilibrium versus punishing with the inefficient static equilibrium whose payoffs are given by u_0. Now α as well as the rest of the proof can be derived as in Rahman (2013, Proposition 3) in the environment of Example 3. The general case with public and private monitoring is dealt in the same way as in Theorem 2, by amending proofs just as in Propositions 8 and 10 if necessary. □