Perfect Versus Imperfect Monitoring in Repeated Games

Takuo Sugaya and Alexander Wolitzky

Stanford Graduate School of Business and Stanford University

December 10, 2013

Abstract

This note presents three results on repeated games. First, players can be better off with imperfect private monitoring than with perfect monitoring, even in the presence of a mediator who can condition her recommendations on the entire history of actions and recommendations. Second, the folk theorem holds with mediated perfect monitoring without any full-dimensionality conditions, so private monitoring cannot improve on mediated perfect monitoring when players are patient. Third, if the mediator can condition her recommendations on actions only, then even patient players can benefit from private monitoring.

Preliminary. Comments Welcome.

1 Introduction

Intertemporal incentives can only be provided when actions are monitored. The purpose of this note is to examine whether it is always the case that the strongest possible intertemporal incentives can provided when monitoring is *perfect*, or alternatively if individuals can sometimes benefit from imperfections in monitoring. While our contribution is purely conceptual, this question seems potentially relevant for understanding the role of information-sharing institutions in settings where repeated game modeling is prevalent, such as long-distance trade (Milgrom, North, and Weingast, 1990; Greif, Milgrom, and Weingast, 1994), collusion (Stigler, 1964; Harrington and Skrzypacz, 2011), organizational economics (Baker, Gibbons, and Murphy, 1994; Levin, 2003), and contemporary international trade (Maggi, 1999; Bagwell and Staiger, 2004). In particular, we argue that there should be no general theoretical presumption that perfect monitoring is preferable to imperfect monitoing for proving incentives in repeated games settings like these.

An obvious way in which imperfect monitoring can help players in a repeated game is that it can give them a way to correlate their actions: if players 1 and 2 observe a signal that player 3 does not observe, they can threaten to hold player 3 below her usual one-shot minmax payoff, and can thus give her stronger incentives than is possible with perfect monitoring (Fudenberg and Tirole, 1991; Gossner and Hörner, 2010). We rule out this effect by assuming that with perfect monitoring the players have access to a *mediator*, who stands in for the players' ability to observe the outcomes of arbitrarily complicated correlating devices (Forges, 1986; Myerson, 1986).[1]

As it turns out, a critical distinction is whether we allow a *general dynamic mediator*, who can condition recommendations on the full history of actions and recommendations, or a *Markov mediator*, who can condition recommendations only on actions (the Markov mediator is so called because it can be replaced each period by a new mediator who knows only the publicly observable history of actions). The key difference between the two kinds of mediator is that with a dynamic mediator, players' continuation play may fail to be common knowledge, which breaks the recursive structure of perfect monitoring repeated games, whereas with a Markov mediator this recursive structure is preserved. In our view, allowing a dynamic mediator may be appropriate if the players can design an information-sharing institution to help them play the game, while a Markov mediator is a more appropriate stand-in for "naturally occurring" correlating devices, such as the phases of the moon used to coordinate bidding in the "electrical conspiracy" documented by Smith (1961).

We present three results. First, players can do better with imperfect private monitoring

[1] With imperfect monitoring, a mediator can aggregate information in addition to correlating play. As we only allow a mediator with perfect monitoring, this issue does not arise here.

than with perfect monitoring, even with a dynamic mediator. Second, the folk theorem holds with perfect monitoring and a dynamic mediator without any full-dimensionality conditions, so *patient* players cannot do (much) better with imperfect monitoring than with perfect monitoring and a dynamic mediator (and the rate of convergence is $1-\delta$, so players need only be moderately patient). Third, even patient players can do better with imperfect monitoring than with perfect monitoring and a Markov mediator. Taken together, our results show that there should be no general presumption that perfect monitoring outperforms imperfect private monitoring in repeated games.

One way of interpreting our results is in the light of previous attempts to establish a recursive upper bound on the equilibrium payoff set in repeated games with private monitoring by adding opportunities for communication or correlation to these games (Renault and Tomala, 2004; Tomala, 2009; Cherry and Smith, 2011). Our first result implies that even the equilibrium payoff set with perfect monitoring and dynamic mediation does not provide such a bound in general. However, a tighter bound may be possible for patient players or for specific classes of games.

Finally, we note that examples by Kandori (1991), Sekiguchi (2002), and Mailath, Matthews, and Sekiguchi (2002) show that players can benefit from imperfect monitoring in finitely repeated games. However, in their examples this conclusion relies on the absence of a dynamic mediator, and is thus ultimately traceable to the possibilities for correlation opened up by private monitoring. We illustrate this in Section 6 below. The broader point that players can benefit from "less" monitoring of each other's actions in games has also been made in various settings, including by Myerson (1991, Section 6.7), Bagwell (1995), Kandori and Obara (2006), and Fuchs (2007); the basic idea that information can be bad for incentives has been known at least since Hirshleifer (1971).

2 Model

The stage game is $G = \left(I, (A_i)_{i\in I}, (u_i)_{i\in I}\right)$, where $I = \{1, \ldots, n\}$ is the set of players, A_i is the finite set of player i's actions, and $u_i : A \to \mathbb{R}$ is player i's payoff function. Players have common discount factor δ. The solution concept is sequential equilibrium. We compare the following information structures.

2.1 Mediated Perfect Monitoring

In period $t = 1, 2, \ldots$, the game proceeds as follows.

1. Given the mediator's history $h_m^t = (m_\tau, a_\tau)_{\tau=1}^{t-1}$ with $h_m^1 = \{\emptyset\}$, the mediator sends a private message $m_{i,t} \in M_i$ to each player i, where M_i is an arbitrary finite set. Let H_m^t be the set of the mediator's period t histories.

2. Given player i's history $h_i^t = \left((m_{i,\tau}, a_\tau)_{\tau=1}^{t-1}, m_{i,t}\right)$ with $h_i^1 = m_{i,1}$, player i takes an action $a_{i,t} \in A_i$. All players and the mediator observe action profile $a_t \in A$. Let H_i^t be the set of player i's period t histories.

A strategy of the mediator's is a map $\sigma_m : \bigcup_{t=1}^{\infty} H_m^t \to \Delta(M)$. The mediator is *Markov* if $\sigma_m(H_m^t)$ depends only on $(a_\tau)_{\tau=1}^{t-1}$ for all H_m^t. A strategy of player i's is a map $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$.

Note that the definition of sequential equilibrium depends on whether we treat the mediator as a "machine" who cannot tremble or as a "player" who can. In the "machine" interpretation, sequential rationality is imposed (and beliefs are defined) only at histories consistent with the mediator's strategy. In the "player" interpretation, sequential rationality is imposed everywhere, and the mediator is treated like any other player in the definition of sequential equilibrium. We adopt the "machine" interpretation throughout, but all our results also hold with the "player" interpretation: the only change would be a slight alteration to the proof of Theorem 1.

2.2 Private Monitoring

In period $t = 1, 2, \ldots$, the game proceeds as follows. Given player i's history $h_i^t = (a_{i,\tau}, z_{i,\tau})_{\tau=1}^{t-1}$ with $h_i^1 = \{\emptyset\}$, player i takes an action $a_{i,t} \in A_i$. Signal $z_t = (z_{i,t})_{i \in I} \in Z$ is drawn from distribution $p(z|a)$. Player i observes $z_{i,t}$. Again, H_i^t is the set of player i's period t histories, and a strategy of player i's is a map $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$.

Note that player i does not observe her payoffs each period. The standard interpretation is that δ represents the probability of the game's continuing, and that payoffs are received when the game ends. This assumption is made purely for ease of exposition. In particular, both of our private monitoring examples may be modified so that players receive payoffs each period.

3 Private Monitoring Versus Perfect Monitoring with Dynamic Mediation

Our first example shows that private monitoring (without mediation) can outperform perfect monitoring with a dynamic mediator.

Player 1 (row player) and player 2 (column player) play a repeated game with common discount factor $\delta = \frac{1}{6}$. The stage game is as follows.

	L	M	R
U	$2,2$	$-1,0$	$-1,0$
D	$3,0$	$0,0$	$0,0$
T	$0,3$	$6,-3$	$-6,-3$
B	$0,-3$	$0,3$	$0,3$

We show that with perfect monitoring with a dynamic mediator there is no sequential equilibrium where the players' per-period payoffs sum to more than 3, while there can be such a sequential equilibrium with private monitoring.[2] The idea is that player 1 can be induced to play U in response to L only if action profile (U, L) is immediately followed by (T, M) with high probability. With perfect monitoring, player 2 always "sees (T, M) coming" after (U, L), and will therefore deviate to L. With private monitoring, player 2 may not know whether (U, L) has just occurred, and therefore may be unsure of whether the next action profile will be (T, M) or (B, M), which gives him the necessary incentive to play M.

[2] In fact, the same is true for Nash equilibrium. This is a second point of contrast with the examples of Kandori (1991), Sekiguchi (2002), and Mailath, Matthews, and Sekiguchi (2002), which work only for sequential equilibrium.

3.1 Perfect Monitoring with Dynamic Mediation

As the players' stage game payoffs from any profile other than (U, L) sum to at most 3, it follows that the players' per-period payoffs may sum to more than 3 only if (U, L) is played in some period t with positive probability. For this to occur in equilibrium, player 1's expected continuation payoff from playing U must exceed her expected continuation payoff from playing D by more than 1, her instantaneous gain from playing D rather than U. Now, player 1 can guarantee herself a continuation payoff of 0 by always playing D, so her expected continuation payoff from playing U must exceed 1. This is possible only if the probability that (T, M) is played in period $t+1$ when U is played in period t exceeds the number p such that

$$\frac{1}{6}\left[p\left(6\right)+\left(1-p\right)\left(3\right)\right]+\frac{1}{6}\underbrace{\left(\frac{1}{6}+\frac{1}{6^2}+\cdots\right)}_{\frac{1}{5}}\left(6\right)=1,$$

or

$$p=\frac{3}{5}.$$

In particular, there must exist a period $t+1$ history h_2^{t+1} of player 2's such that (T, M) is played with probability at least $\frac{3}{5}$ in period 1 conditional on reaching h_2^{t+1}. But, at such a history, player 2's payoff from playing M is at most

$$\frac{3}{5}\left(-3\right)+\frac{2}{5}\left(3\right)+\frac{1}{5}\left(3\right)=0,$$

while, noting that player 2 can guarantee himself continuation payoff 0 by playing $\frac{1}{2}L+\frac{1}{2}M$, player 2's payoff from playing L is at least

$$\frac{3}{5}\left(3\right)+\frac{2}{5}\left(-3\right)+\frac{1}{5}\left(0\right)=\frac{3}{5}.$$

Therefore, player 2 will deviate at such a history, so no such equilibrium can exist.

3.2 Private Monitoring

Consider the following imperfect private monitoring structure. Player 1 always observes an uninformative signal $\emptyset$. There are two possible private signals for player 2, m and r. Conditional on any action profile in which player 1 plays U, signal m obtains with probability 1. Conditional on any other action profile, signals m and r obtain with probability $\frac{1}{2}$ each.

We now describe a strategy profile under which the players' payoffs sum to $\frac{23}{7} \approx 3.29$.

Player 1's strategy: In each odd period $t = 2n + 1$ with $n = 0, 1, ...$, player 1 plays $\frac{1}{3}U + \frac{2}{3}D$. Let $a_1(n)$ be the realization of this mixture. In the even period $t = 2n + 2$, if the previous action $a_1(n) = U$, then player 1 plays T; if the previous action $a_1(n) = D$, then player 1 plays B.

Player 2's strategy: In each odd period $t = 2n + 1$ with $n = 0, 1, ...$, player 2 plays L. Let $y_2(n)$ be the realization of player 2's private signal. In the even period $t = 2n + 2$, if the previous private signal $y_2(n) = m$, then player 2 plays M; if the previous signal $y_2(n) = r$, then player 2 plays R.

We check that this strategy profile, together with any consistent belief system, is a sequential equilibrium.

In an odd period, player 1's payoff from U is the solution to

$$v = 2 + \frac{1}{6}(6) + \frac{1}{6^2}v.$$

On the other hand, her payoff from D is

$$3 + \frac{1}{6}(0) + \frac{1}{6^2}v.$$

Hence, player 1 is indifferent between U and D (and clearly prefers either of these actions to T or B).

In addition, playing L is a myopic best response for player 2, player 1's continuation play is independent of player 2's action, and the distribution of player 2's signal is independent of player 2's action. Hence, playing L is optimal for player 2.

Next, in an even period, it suffices to check that both players always play myopic best responses, as in even periods continuation play is independent of realized actions and signals. If player 1's last action was $a_1(n) = U$, then she believes that player 2's signal is $y_2(n) = m$ with probability 1 and thus that he will play M. Hence, playing T is optimal. If instead player 1's last action was $a_1(n) = D$, then she believes that player 2's signal is equal to m and r with probability $\frac{1}{2}$ each, and thus that he will play $\frac{1}{2}M + \frac{1}{2}R$. Hence, both T and B are optimal.

On the other hand, if player 2 observes signal $y_2(n) = m$, then his posterior belief that player 1's last action $a_1(n) = U$ is

$$\frac{\frac{1}{3}(1)}{\frac{1}{3}(1) + \frac{2}{3}\left(\frac{1}{2}\right)} = \frac{1}{2}.$$

Hence, player 2 is indifferent among all of his actions. If player 2 observes $y_2(n) = r$, then his posterior is that $a_1(n) = D$ with probability 1, so that M and R are optimal.[3]

Finally, expected payoffs under this strategy profile in odd periods sum to

$$\underbrace{\frac{1}{3}(4)}_{\text{player 1 plays } U} + \underbrace{\frac{2}{3}(3)}_{\text{player 1 plays } D} = \frac{10}{3},$$

and in even periods sum to

$$\underbrace{\frac{1}{3}(3)}_{\Pr(T,M)(6-3)} + \underbrace{\frac{1}{3}(3)}_{\Pr(B,M)(3)} + \underbrace{\frac{1}{3}(3)}_{\Pr(B,R)(3)} = \frac{9}{3}.$$

Therefore, per-period expected payoffs sum to

$$\left(1 - \frac{1}{6}\right)\left(\frac{10}{3} + \frac{1}{6}\left(\frac{9}{3}\right)\right)\left(1 + \frac{1}{6^2} + \frac{1}{6^4} + \dots\right) = \frac{23}{7}.$$

4 The Folk Theorem with Perfect Monitoring and Dynamic Mediation

We now show that the folk theorem always holds with perfect monitoring and dynamic mediation, with a rate of convergence of $1 - \delta$. Therefore, in contrast to our first example, moderately patient players cannot significantly benefit from imperfect monitoring in the presence of a dynamic mediator.

Fix a mediated perfect monitoring game. Let $\underline{u}_i$ be player i's correlated minmax payoff, given by

$$\underline{u}_i = \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i\left(a_i, \alpha_{-i}\right).$$

Let d_i be the greatest possible difference in player i's payoff resulting from a change in player i's action only, so that

$$d_i = \max_{a \in A, a_i' \in A_i} u_i\left(a_i', a_{-i}\right) - u_i\left(a\right).$$

[3] These indifferences result because we have chosen payoffs to make the example as simple as possible. The example is generic.

Let $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$ and $d = (d_1, \ldots, d_n)$. Recall that a payoff vector $v \in \mathbb{R}^n$ is *feasible* if $v = u(\alpha)$ for some $\alpha \in \Delta(A)$. Let $F \subseteq \mathbb{R}^n$ denote the set of feasible payoff vectors, and let $\mathring{F}$ denote the interior of this set. Say that v is *strictly individually rational* if $v > \underline{u}$. Note that this is more permissive than the usual notion of strict individual rationality as it is defined relative to correlated minmax payoffs. Let $E(\delta)$ denote the sequential equilibrium payoff set.

Our result is the following.

Theorem 1 *If a payoff vector $v \in \mathring{F}$ satisfies*

$$v \geq \underline{u} + \frac{1-\delta}{\delta} d, \tag{1}$$

then v is a sequential equilibrium payoff vector. That is,

$$\left\{ v \in \mathring{F} : v \geq \underline{u} + \frac{1-\delta}{\delta} d \right\} \subseteq E(\delta).$$

With the caveat that it does not address the attainability of payoff vectors on the boundary of F, Theorem 1 is substantially stronger than the folk theorem, as it explicitly gives a discount factor sufficient for each feasible and strictly individually rational payoff vector to be attainable in sequential equilibrium. In particular, the rate of convergence of $E(\delta)$ to F is $1-\delta$. We nonetheless record the folk theorem as a corollary.

Corollary 1 (Folk Theorem) *For every strictly individually rational payoff vector $v \in \mathring{F}$, there exists $\bar{\delta} < 1$ such if $\delta > \bar{\delta}$ then $v \in E(\delta)$.*

Proof. Take $\bar{\delta} = \max_{i \in I} \frac{d_i}{v_i - \underline{u}_i + d_i}$ and apply Theorem 1. ■

The intuition behind Theorem 1 is that the mediator can recommend mixed actions leading to the target payoff v with high probability, while recommending every action profile with positive probability. Following a deviation by player i, the mediator can recommend that i's opponents minmax her forever, without alerting them to the fact that a deviation occurred (in particular, i's opponents believe that the "deviation" was in fact recommended by the mediator, since the recommendation has full support). The key point is then that i's opponents are willing to minmax i forever, even though this may be very costly for them, because they never realize that a deviation has occurred and always expect to return to getting payoff v in the next period.

Proof of Theorem 1. Fix a payoff vector $v \in \mathring{F}$ satisfying (1). Let $\bar{v} = \frac{1}{|A|}\sum_{a \in A} u(a)$ be the payoff vector that results when the players mix uniformly over all possible pure action profiles. Choose $\varepsilon > 0$ small enough such that the payoff vector v^* given by

$$v^* = \frac{v - \varepsilon \bar{v}}{1 - \varepsilon}$$

is feasible; this is possible by the assumption that $v \in \mathring{F}$. Let $\alpha^* \in \Delta(A)$ be such that $u(\alpha^*) = v^*$.

Consider the following mediator's strategy. There are $n+1$ states, a regular state R and a punishment state for each player i, P_i. Begin in state R.

- In state R, recommend α^* with probability $1 - \varepsilon$, and recommend each pure action profile with probability $\frac{\varepsilon}{|A|}$ (that is, recommend $(1-\varepsilon)\alpha^* + \sum_{a \in A} \frac{\varepsilon}{|A|} a$). If exactly one player i does not follow her recommendation, go to state P_i. Otherwise, stay in state R.

- In state P_i, recommend $\hat{\alpha}_{-i}$ to players $-i$, for some

  $$\hat{\alpha}_{-i} \in \arg \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}),$$

  and recommend $\hat{a}_i$ to player i, for some

  $$\hat{a}_i \in \max_{a_i \in A_i} u_i(a_i, \hat{\alpha}_{-i}).$$

  Stay in state P_i.

We check that it is optimal for players to obey the mediator's recommendations.

We first claim that after any history of actions and private recommendations, each player i either assigns probability 1 to the state being R or assigns probability 1 to the state being P_i. More specifically, we claim that if player i has disobeyed the mediator more recently than she last received a recommendation $a_i \neq \hat{a}_i$, then she is certain that the state is P_i; while if player i received a recommendation $a_i \neq \hat{a}_i$ more recently than she has disobeyed the mediator (or if she has never disobeyed the mediator), then she is certain that the state is R. To see this, first note that if player i has never disobeyed the mediator then she is certain the state is R, as all possible histories of recommendations and opposing actions are on-path. Starting at a history h_i^t where player i is certain the state is R, if player i disobeys the mediator and is recommended $\hat{a}_i$ in every period $\tau = t+1, \ldots, t'$, then player i is certain the state is P_i at history $h_i^{t'}$. Starting from a history h_i^t where player i is certain the state is P_i, if player i is recommended $a_i \neq \hat{a}_i$ then player i learns that in fact that state has never been P_i (as once the mediator enters state P_i she recommends $\hat{a}_i$ forever, and the mediator never trembles). This can only be because another player also disobeyed the mediator in every period where player i disobeyed the mediator, and since trembles are independent across information sets this implies that the current state is R with probability 1. Thus, we have shown that player i begins the game certain that the state is R, switches from being certain that the state is R to being certain that the state is P_i whenever she disobeys the mediator, and switches back to being certain that the state is R whenever she is recommended an action $a_i \neq \hat{a}_i$. This proves the claim.

It remains only to check that it is optimal for each player i to obey the mediator when she is certain the state is R and when she is certain the state is P_i.

If player i is certain the state is R, her instantaneous gain from disobeying the current recommendation is at most $(1-\delta)\, d_i$, while her loss in continuation payoff is

$$\delta\left(\left((1-\varepsilon)\, v^* + \varepsilon \bar{v}\right) - \underline{u}_i\right) = \delta\left(v - \underline{u}_i\right).$$

It follows from (1) that

$$(1-\delta)\, d_i \leq \delta\left(v - \underline{u}_i\right),$$

so obeying the current recommendation is optimal.

If player i is certain the state is P_i, then as we have seen she must have been recommended $\hat{a}_i$ (as at any history, receiving recommendation $a_i \neq \hat{a}_i$ convinces her that the state is R). Thus, obeying the mediator is a myopic best response, and her continuation payoff is independent of her current action, so obeying is optimal. ∎

5 Private Monitoring Versus Perfect Monitoring with Markov Mediation

It may be sometimes more natural to assume that the mediator (if available) can condition her recommendations only on actions and not on past recommendations. For example, this will be the case if there is a different mediator every period, and the period t mediator knows the publicly observable history of actions but not the private messages sent by the period $t-1$ mediator. More abstractly, the key feature of Markov mediation is that it preserves common knowledge of continuation play, and thus preserves the recursive structure of the repeated game.[4]

This section shows that restricting to Markov mediation is not without loss of generality. In particular, with Markov mediation even patient players can benefit from private monitoring.

To see this, consider the following slight modification of a well-known example due to Fudenberg and Maskin (1986). There are four players. Player 1 picks a row, player 2 picks a column, and player 3 picks a matrix. Player 4 is a dummy player. The payoff matrix is as follows.

	l	r		l	r
T	$1,1,1,0$	$0,0,0,100$	T	$0,0,0,100$	$0,0,0,100$
B	$0,0,0,100$	$0,0,0,100$	B	$0,0,0,100$	$1,1,1,0$
	L			R	

Note that player 4's payoff is high when the payoff of players 1, 2, and 3 is low, and vice versa.

5.1 Perfect monitoring with Markov Mediation

Fudenberg and Maskin (1986) show that, without mediation, the equilibrium payoff of players 1, 2, and 3 cannot be less than $\frac{1}{4}$. Their proof extends easily to the case with a Markov mediator.

[4] With perfect monitoring, sequential equilibrium with a Markov mediator corresponds to Tomala's (2009) perfect communication equilibrium.

For any mixed action $\alpha \in \Delta(A)$, let $\alpha_i(1)$ be the probability attached to the first action for player $i \in \{1,2,3\}$ (that is, to T, l, or L; we neglect the dummy player 4). Then there must be a player i who faces $\alpha_j(1) \geq \frac{1}{2}$ and $\alpha_k(1) \geq \frac{1}{2}$, or $\alpha_j(1) \leq \frac{1}{2}$ and $\alpha_k(1) \leq \frac{1}{2}$, for $j \neq i$ and $k \neq i, j$. In the former case, playing the first action gives player i a payoff no less than $\frac{1}{4}$, while in the latter, playing the second action does so.

With a Markov mediator, the distribution of future paths of play depends only on the public history of actions $(a_\tau)_{\tau=1}^{t-1}$. Let $\underline{v} \in \mathbb{R}$ be the infimum over all histories of actions and all sequential equilibria of the (common) continuation payoff of players 1, 2, and 3. As the current recommendation plan is common knowledge, at every history there is at least one player $i \in \{1,2,3\}$ who can guarantee an instantaneous payoff of $\frac{1}{4}$. Hence,

$$\underline{v} \geq (1-\delta)\frac{1}{4} + \delta\underline{v}.$$

That is, $\underline{v} \geq \frac{1}{4}$.

It follows that player 4's equilibrium payoff cannot exceed $\frac{300}{4}$. Therefore, the sum of all the players' payoffs is no more than

$$1 + 1 + 1 + \frac{300}{4} = 78.$$

5.2 Private monitoring

We exhibit an information structure for which the sum of the players' payoffs may be arbitrarily close to 100.

Each player $i \in \{1,2,3\}$ observes two signals, $y_i \in \{A, B, O\}$ and $z_i \in \{0,1\}$. Player 4 observes nothing.

The distribution of signal y is determined by player 3's action, as follows.

1. If $a_3 = L$, then with probability ε^2, all players observe $y_i = A$; with probability ε^2, all players observe $y_i = B$; and with probability $1 - 2\varepsilon^2$, all players observe $y_i = O$.

2. If $a_3 = R$, then with probability ε^2, player 1 observes A, player 2 observes B, and player 3 observes O; with probability ε^2, player 1 observes B, player 2 observes A, and player 3 observes O; and with probability $1 - 2\varepsilon^2$, all players observe $y_i = O$.

Note that if player 3 plays R then players 1 and 2's signals may be "miscoordinated."

The distribution of signal z is given by:

1. If $a_1 = T$ and $a_2 = l$, or if $a_1 = B$ and $a_2 = r$, then all players observe $z_i = 1$ with probability ε, and all players observe $z_i = 0$ with probability $1 - \varepsilon$.

2. Otherwise, all players observe $z_i = 0$ with probability 1.

Note that players 1 and 2 must "coordinate" in order to generate a public signal $z_1 = z_2 = z_3 = 1$.

For each $\varepsilon > 0$, we construct an equilibrium with payoffs converging to $\frac{2\varepsilon}{1+2\varepsilon}$ for players 1, 2, and 3 as $\delta \to 1$. By feasibility, player 4's payoff converges to 100. We ignore player 4 throughout the construction.

Each player $i \in \{1, 2, 3\}$ has state $s_{i,t} \in \{A, B, O\}$ in period t. The initial state is $s_{1,1} = s_{2,1} = s_{3,1} = O$. For each period t,

1. If $s_{i,t} = O$, then player $i = 1$ plays B, player $i = 2$ plays r, and player $i = 3$ plays L. If $y_{i,t} = A$, go to state $s_{i,t+1} = A$. If $y_{i,t} = B$, go to state $s_{i,t+1} = B$. If $y_{i,t} = O$, go to state $s_{i,t+1} = O$.

2. If $s_{i,t} = A$, then player $i = 1$ plays T, player $i = 2$ plays l, and player $i = 3$ plays L. If $z_{i,t} = 0$, go to state $s_{i,t+1} = A$. If $z_{i,t} = 1$, go to state $s_{i,t+1} = O$.

3. If $s_{i,t} = B$, then player $i = 1$ plays B, player $i = 2$ plays r, and player $i = 3$ plays R. If $z_{i,t} = 0$, go to state $s_{i,t+1} = B$. If $z_{i,t} = 1$, go to state $s_{i,t+1} = O$.

Intuitively, if player 3 deviates and plays R when $s_{1,t} = s_{2,t} = s_{3,t} = O$, then the next state might be $(s_{1,t+1}, s_{2,t+1}) = (A, B)$ or (B, A). If the state reaches $(s_{1,t+1}, s_{2,t+1}) = (A, B)$ or (B, A), then players 1 and 2 miscoordinate forever, yielding payoff 0 for players 1, 2, and 3. Note that this "punishment" for player 3's deviation is like the punishments underlying Theorem 1: players 1 and 2 never learn that a deviation has occurred, so they minmax player 3 (and themselves) forever, while always believing that they are still on the equilibrium path.

We check that this strategy profile, together with any consistent belief system, is a sequential equilibrium.

Let $v_i(s_i)$ be player i's equilibrium continuation value in state s_i, given by

$$\begin{aligned} v_i(O) &= (1-\delta)(0) + \delta\left(\left(1-2\varepsilon^2\right)v_i(O) + \varepsilon^2 v_i(A) + \varepsilon^2 v_i(B)\right), \\ v_i(A) &= (1-\delta)(1) + \delta\left((1-\varepsilon)v_i(A) + \varepsilon v_i(O)\right), \\ v_i(B) &= (1-\delta)(1) + \delta\left((1-\varepsilon)v_i(B) + \varepsilon v_i(O)\right). \end{aligned}$$

Solving this system of equations yields

$$\begin{aligned} v_i(O) &= \frac{2\delta\varepsilon^2}{1-\delta+\delta\varepsilon+2\delta\varepsilon^2}. \\ v_i(A) &= v_i(B) = \frac{1-\delta+2\delta\varepsilon^2}{1-\delta+\delta\varepsilon+2\delta\varepsilon^2}. \end{aligned}$$

It is straightforward to check that players 1 and 2 do not have an incentive to deviate: In state $s_{i,t} = O$, player 1 or 2 cannot control the instantaneous utility or transition probability. In state $s_{i,t} = A$ or B, conforming gives payoff $v_i(s_{i,t})$, while deviating gives payoff $\delta v_i(s_{i,t})$ as it gives instantaneous utility 0.

For player 3, the only problematic state is $s_{3,t} = O$, as in the other states she cannot control the instantaneous utility or transition probability. When $s_{3,t} = 0$, we can compare player 3's continuation payoff from L and R as follows.

1. With probability $1-2\varepsilon^2$, the next state is $s_{1,t} = s_{2,t} = s_{3,t} = O$ regardless of a_3, so player 3's continuation payoff is independent of a_3.

2. With probability $2\varepsilon^2$,

    (a) With L, the next state is $s_{1,t+1} = s_{2,t+1} = s_{3,t+1} = A$ or $s_{1,t+1} = s_{2,t+1} = s_{3,t+1} = B$, so player 3's continuation payoff is $v_3(A) = v_3(B) = \frac{1-\delta+2\delta\varepsilon^2}{1-\delta+\delta\varepsilon+2\delta\varepsilon^2}$.

    (b) With R, players 1 and 2 will be permanently miscoordinated: $(s_{1,\tau}, s_{2,\tau}) = (A, B)$ or (B, A) for all $\tau > t$. Player 3's continuation payoff is 0.

Hence, player 3 will conform if

$$1-\delta \le 2\varepsilon^2 \frac{1-\delta+2\delta\varepsilon^2}{1-\delta+\delta\varepsilon+2\delta\varepsilon^2}.$$

As the left hand side converges to 0 while the right hand side converges to $\frac{4\varepsilon^3}{1+2\varepsilon}$, player 3 will conform for high enough δ. Finally, the players' equilibrium payoff is

$$v_i(O) = \frac{2\delta\varepsilon^2}{1-\delta+\delta\varepsilon+2\delta\varepsilon^2} \rightarrow_{\delta\rightarrow 1} \frac{2\varepsilon}{1+2\varepsilon},$$

as desired.

5.3 Scope of the Example

In the above example, the sum of the players' equilibrium payoffs is higher with private monitoring than with perfect monitoring and Markov mediation. A natural question is whether private monitoring can strictly Pareto dominate perfect monitoring with Markov mediation. That is, letting $E_{\text{Markov}}(\delta)$ be the equilibrium payoff set with perfect monitoring and Markov mediation, and letting $E_{\text{Markov}} = \lim_{\delta\rightarrow 1} E_{\text{Markov}}(\delta)$ be the limit equilibrium payoff set, can we find for every δ a monitoring structure and equilibrium such that the equilibrium payoff $v(\delta)$ converges to v such that $v_i > w_i$ for all $i \in I$ and $w \in E_{\text{Markov}}$?

The answer is no. Take a feasible payoff vector v that strictly Pareto dominates $w \in E_{\text{Markov}}(\delta)$ for some $\delta < 1$. Since v is feasible, for sufficiently large δ, there exists a sequence of action profiles $(a_\tau)_{\tau=1}^{T}$ with $T < \infty$ such that the average payoff from the sequence is equal to v: $\frac{1-\delta}{1-\delta^T}\sum_{\tau=1}^{T}\delta^{\tau-1}u(a_\tau) = v$.[5] In addition, $E_{\text{Markov}}(\delta)$ is monotone by standard arguments, so $w \in E_{\text{Markov}}(\delta')$ for all $\delta' \geq \delta$. Then, for sufficiently large δ, the following strategy profile is an equilibrium: play a_t in period $t \pmod T$ unless there has been a unilateral deviation; after a unilateral deviation, play the equilibrium whose payoff is w.

It is also natural to look for sufficient conditions for private monitoring not to outperform perfect monitoring with Markov mediation. If there are only two players, or if the NEU condition is satisfied as in Abreu, Dutta, and Smith (1994), then the correlated minmax folk theorem holds with perfect monitoring and Markov mediation: $E_{\text{Markov}} = \{v \in F : v \geq \underline{u}\}$. Under these conditions, the limit equilibrium payoff set with private monitoring is a subset of E_{Markov}.

[5] The proof is the same as Fudenberg and Maskin's (1991) proof of the dispensability of public randomization, and so is omitted.

6 Comparison with Kandori (1991), Sekiguchi (2002), and Mailath, Matthews, and Sekiguchi (2002)

Kandori (1991), Sekiguchi (2002), and Mailath, Matthews, and Sekiguchi (2002) present examples in which players benefit from imperfect monitoring. This section shows that in Kandori's and Sekiguchi's examples this conclusion relies on the absence of a dynamic mediator. The same is true of Mailath, Matthews, and Sekiguchi's examples, for similar reasons.

Let $NE(G)$ denote the set of stage game Nash equilibria of G. Condition 1 of Sekiguchi (2002), which generalizes Kandori (1991), is as follows.

Condition 1 *There exists $a^* \in A \setminus NE(G)$, $\alpha^* \in NE(G)$, and $\alpha^p \in \Delta(A)$ such that*

1. *For each i, the support of α_i^p is a subset of the support of α_i^*.*

2. *For any i,*

$$u_i(\alpha^*) - \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}^p) > \max_{a_i \in A_i} u_i(a_i, a_{-i}^*) - u_i(a^*).$$

Sekiguchi shows that in the undiscounted two-period repetition of G, there is a private monitoring structure and a corresponding equilibrium in which a^* is played in period 1 and α^* is played in period 2. By an argument similar to the proof of Theorem 1, the same is true with perfect monitoring and dynamic mediation.

In particular, let

$$BR_i^\varepsilon = \{a_i \in A_i : u_i(a_i, a_{-i}^*) + \varepsilon \geq u_i(a^*)\}.$$

Consider the following mediator's strategy.

- In period 1, recommend a^* with probability $1 - \eta$ and recommend $\left(a_i, a_{-i}^*\right)$ with probability $\frac{\eta}{|I|\left|BR_i^\varepsilon\right|}$ for each $a_i \in BR_i^\varepsilon$ and $i \in I$.

- In period 2, if there was a unilateral deviation by player i to $a_i \in BR_i^\varepsilon$ in period 1, recommend $\left(BR_i\left(\alpha_{-i}^p\right), \alpha_{-i}^p\right)$. Otherwise, recommend α^*.

Take $\varepsilon > 0$ and $\eta > 0$ sufficiently small so that, for all $i \in I$,

$$u_i(\alpha^*) - \max_{a_i \in A_i} u_i(a_i, \alpha^p_{-i}) > \max_{a_i \in A_i} u_i(a_i, a^*_{-i}) - u_i(a^*) + 2\eta \max_a |u_i(a)| + \varepsilon, \tag{2}$$

$$\varepsilon > \frac{2\eta}{1-\eta} \max_a |u_i(a)|. \tag{3}$$

It is straightforward to see that each player has an incentive to follow the recommendation in period 2. Consider period 1. When a^*_i is recommended, player i's payoff from following the recommendation is at least

$$(1-\eta) u_i(a^*) - \eta \max_a |u_i(a)| + u_i(\alpha^*).$$

When $a_i \in BR^\varepsilon_i$ is recommended, player i's payoff from following the recommendation is at least

$$u_i(a^*) - \varepsilon + u_i(\alpha^*).$$

In either case, player i's payoff from deviating to $a'_i \in BR^\varepsilon_i$ is at most

$$(1-\eta) \max_{a_i \in A_i} u_i(a_i, a^*_{-i}) + \eta \max_a |u_i(a)| + \max_{a_i \in A_i} u_i(a_i, \alpha^p_{-i})$$

Hence, in either case, (2) implies that deviating to $a_i \in BR^\varepsilon_i$ is unprofitable. Finally, player i's payoff from deviating to $a'_i \notin BR^\varepsilon_i$ when a^*_i is recommended is at most

$$(1-\eta)(u_i(a^*) - \varepsilon) + \eta \max_a |u_i(a)| + u_i(\alpha^*),$$

while her payoff from deviating to $a'_i \notin BR^\varepsilon_i$ when $a_i \in BR^\varepsilon_i$ is recommended is at most

$$u_i(a^*) - \varepsilon + u_i(\alpha^*).$$

Hence, in the first case (3) implies that deviating to $a'_i \notin BR^\varepsilon_i$ is unprofitable, while in the second case such a deviation is clearly unprofitable.

References

[1] Bagwell, K. (1995), "Commitment and Observability in Games," *Games and Economic Behavior*, 8, 271-280.

[2] Bagwell, K.W. and R.W. Staiger (2004), *The Economics of the World Trading System*, MIT Press: Cambridge.

[3] Baker, G., R. Gibbons, and K.J. Murphy (2002), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 117, 39-84.

[4] Cherry, J. and L. Smith (2011), "Unattainable Payoffs for Repeated Games of Private Monitoring," *mimeo.*

[5] Forges, F. (1986), "An Approach to Communication Equilibria," *Econometrica*, 54, 1375-1385.

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

[7] Fudenberg, D. and E. Maskin (1991), "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, 53, 428-438.

[8] Fuchs, W. (2007), "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review*, 97, 1432-1448.

[9] Gossner, O. and J. Hörner (2010), "When is the Lowest Equilibrium Payoff in a Repeated Game Equal to the Minmax Payoff?," *Journal of Economic Theory*, 145, 63-84.

[10] Greif, A., P. Milgrom, and B.R. Weingast (1994), "Coordination, Commitment, and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy*, 102, 745-776.

[11] Harrington, J.E. and A. Skrzypacz (2011), "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices," *American Economic Review*, 101, 2425-2449.

[12] Hirshleifer, J. (1971), "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561-574.

[13] Kandori, M. (1991), "Cooperation in Finitely Repeated Games with Imperfect Private Information," *mimeo.*

[14] Kandori, M. and I. Obara (2006), "Less is More: An Observability Paradox in Repeated Games," *International Journal of Game Theory*, 34, 475-493.

[15] Levin, J. (2003), "Relational Incentive Contracts," *American Economic Review*, 93, 835-857.

[16] Mailath, G.J., S.A. Matthews, and T. Sekiguchi (2002), "Private Strategies in Finitely Repeated Games with Imperfect Public Monitoring," *Contributions in Theoretical Economics*, 2.

[17] Maggi, G. (1999), "The Role of Multilateral Institutions in International Trade Cooperation," *American Economic Review*, 89, 190-214.

[18] Milgrom, P.R. and D.C. North, "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs," *Economics & Politics*, 2, 1-23.

[19] Myerson, R.B. (1986), "Multistage Games with Communication," *Econometrica*, 54, 323-358.

[20] Myerson, R.B. (1991), *Game Theory: Analysis of Conflict*, Harvard University Press: Cambridge.

[21] Renault, J. and T. Tomala (2004), "Communication Equilibrium Payoffs in Repeated Games with Imperfect Monitoring," *Games and Economic Behavior*, 49, 313-344.

[22] Sekiguchi, T. (2002), "Existence of Nontrivial Equilibria in Repeated Games with Imperfect Private Monitoring," *Games and Economic Behavior*, 40, 299-321.

[23] Smith, R.A. (1961), "The Incredible Electrical Conspiracy," *Fortune*, 63, 132-80.

[24] Stigler, G.J. (1964), "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44-61.

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