Bounding Payoffs in Repeated Games with Private Monitoring: n-Player Games

Takuo Sugaya and Alexander Wolitzky

Stanford Graduate School of Business and MIT

December 15, 2016

Abstract

We provide a recursive upper bound on the sequential equilibrium payoff set at a fixed discount factor in repeated games with imperfect private monitoring. The bounding set is the equilibrium payoff set in a repeated "delegation game," where in every period a mediator asks for the players' approval to follow a history-contingent behavior strategy. This set admits a tractable recursive characterization and can thus be applied "off-the-shelf" to bound equilibrium payoffs in any private monitoring repeated game. We illustrate our results with an application to a simple public goods game.

1 Introduction

Repeated games with imperfect private monitoring have been a major topic of research for some time and have been used to model important economic settings such as collusion with secret price cuts (Stigler, 1964) and relational contacting with subjective performance evalutions (Levin, 2003; MacLeod, 2003; Fuchs, 2007). While extensive progress has been made in analyzing the equilibrium set in these games in the limit where players become very patient (see Sugaya, 2016, and references therein), our understanding of equilibria at fixed discount factors—arguably the more relevant case for most economic applications—remains much more limited. In particular, no simple characterization of the sequential equilibrium payoff set at a fixed discount factor is available for repeated games with private monitoring, in contrast to the case of perfect public equilibria in games with public monitoring, where such a characterization is provided by Abreu, Pearce, and Stacchetti (1990; henceforth APS). Kandori (2002) discusses the well-known difficulties involved in generalizing the results of APS to private monitoring.

In a recent paper (Sugaya and Wolitzky, 2016), we have shown that the equilibrium payoff set in a private monitoring repeated game is bounded by the equilibrium set in the same game played under perfect monitoring with the assistance of a mediator, so long as attention is restricted to two-player games where the discount factor exceeds a cutoff value δ^*. This bound is as tractable as APS's characterization for public monitoring games, and it is also tight insofar as mediated perfect monitoring can itself be viewed as a kind of private monitoring. However, the bound does not apply to games with more than two player or when the discount factor is too low.

In this article, we derive a more permissive bound on the equilibrium payoff set in private monitoring repeated games, which however applies to *any* game, regardless of the number of players or (up to a small technicality) the discount factor. The bound again admits a simple, APS-style recursive characterization. Thus, while the bound we provide is not always tight, it gives a simple off-the-shelf method of bounding equilibrium payoffs in any repeated game with private monitoring.

The results of this paper come from considering the repetition of a "delegation game," where in every period the mediator asks for the players' approval to follow a history-contingent behavior strategy. Each player is free to ignore the mediator and take a different action. We first prove that the equilibrium payoff set of this delegation game upper-bounds the equilibrium payoff set of any repeated game with private monitoring. We then offer a recursive characterization of the equilibrium payoff set of the delegation game, which we illustrate in examples.

2 Repeated Games with Private Monitoring

A finite stage game $G = \left(I, (A_i, u_i)_{i \in I}\right)$ is repeated in periods $t = 1, 2, \ldots$, where $I = \{1, \ldots, |I|\}$ 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 maximize expected discounted payoffs with common discount factor $\delta \in (0, 1)$.

In each period t, the game proceeds as follows: Each player i takes an action $a_{i,t} \in A_i$. A signal $y_t = (y_{i,t})_{i \in I} \in \prod_{i \in I} Y_i = Y$ is drawn from distribution $p(y_t|a_t)$, where Y_i is the finite set of player i's signals and (Y, p) (often abbreviated to p) is the monitoring structure. Player i observes $y_{i,t}$.

A period t history of player i's is an element of $H_i^t = (A_i \times Y_i)^{t-1}$, with typical element $h_i^t = (a_{i,\tau}, y_{i,\tau})_{\tau=1}^{t-1}$, where H_i^1 consists of the null history $\emptyset$. A (behavior) strategy of player i's is a mapping $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$. Let $E(\delta, p)$ be the set of Nash equilibrium payoffs in the repeated game with discount factor δ and monitoring structure p.

3 Repeated Delegation to a Mediator

We will also consider a repeated game of delegated action choice to a mediator. In every period of this game, players will have the choice of *approving* the mediator's recommendation or *disapproving* and playing any action in A_i.

Let $C = (C_i)_{i \in I} := (A_i \cup \{approve\})_{i \in I}$ with typical element $c = (c_i)_{i \in I} \in (A_i \cup \{approve\})_{i \in I}$. Given the stage game G, the corresponding *repeated delegation game* proceeds as follows in each period t:

1. The mediator publicly recommends a *mapping* from histories to correlated actions $\mu_t : H_m^t \to \Delta(A)$, where $H_m^t := (M \times C \times A)^{t-1}$ with typical element $h_m^t = (\mu_\tau, c_\tau, a_\tau)_{\tau=1}^{t-1}$ is the mediator's history. Let $H_m^1 = \emptyset$.

2. After observing μ_t, each player i simultaneously chooses an alternative from the set $C_i = A_i \cup \{approve\}$, where the alternative *approve* is interpreted as approving the mediator's recommendation for that player and the alternative $a_i \in A_i$ interpreted as disapproving the mediator's recommendation and instead playing action a_i.

(Note that the mediator does *not* announce the history h_m^t. As we will see, players typically face uncertainty regarding h_m^t. Thus, players must approve or disapprove the mapping μ_t without knowing the resulting correlated action $\mu(h_m^t)$.)

3. If the set of players $J \subseteq I$ approves and the set $-J = I \backslash J$ does not approve and chooses actions a_{-J}, then given the mediator's history h_m^t the realized action profile is drawn from the distribution $(\mu_J(h_m^t), a_{-J})$, where $\mu_J(h_m^t) = \mathrm{marg}_J \mu(h_m^t)$. The realized action profile is perfectly observed. Thus, player i's history at the beginning of period $t+1$ is $h_i^{t+1} := (\mu_\tau, c_{i,\tau}, a_\tau)_{\tau=1}^t$, where for every τ μ_τ is the recommended mapping, $c_{i,\tau} \in A_i \cup \{approve\}$ is player i's chosen alternative, and a_τ is the realized action profile. Let H_i^t denote the set of all histories h_i^t.

We consider sequential equilibria in the repeated delegation game that satisfy the *obedience* condition that players always choose $c_{i,t} = approve$ at all histories, on and off the equilibrium path. Let $E_{del}(\delta)$ be the set of obedient sequential equilibrium payoffs in the repeated delegation game. Note that the "revelation principle" suggests that the restriction to obedient equilibria is without loss of generality. We do not have a proof of this result for this game, but this is irrelevant for our purposes as we will show that $E_{del}(\delta) \supset E(\delta, p)$ for every monitoring structure p even with the restriction to obedient equilibria in the delegation game. Note that considering obedient sequential equilibria in the repeated delegation game and Nash equilibria in the unmediated repeated game only strengthens this result.[1]

Two more details regarding the definition of sequential equilibrium: (i) In specifying an equilibrium, we define beliefs and impose sequential rationality only at histories consistent with the mediator's strategy. (The interpretation is that the mediator is not a player in the game but rather a "machine" that cannot tremble.) (ii) We endow the space of beliefs and strategies with the product topology. Neither of these choices is important for our results.

Note that players do not observe their opponents' choices from C_{-i}. Hence, in an obedient equilibrium, if there exists a realization of $(\mu_\tau)_{\tau=1}^{t-1}$ that results in history h_i^t with positive probability when $c_{j,\tau} = approve$ for all $j \neq i$ and $\tau \leq t-1$, then at history h_i^t player i must

[1] This approach is the same as in Sugaya and Wolitzky (2016). There, we phrased our results as bounding the sequential equilibrium payoff set in unmediated repeated games, but the same results hold for Nash equilibrium.

believe with probability 1 that $c_{j,\tau} = \textit{approve}$ for all $j \neq i$ and $\tau \leq t-1$.

Finally, note that it is crucial that the mediator recommends a mapping *before* players choose whether to approve or deviate. If instead players decided whether to delegate their choice of action to the mediator before the mediator's choice, then any feasible and individually rational payoff vector could be sustained in equilibrium for *any* discount factor, as the mediator could immediately minimax any player who failed to delegate.[2] This alternative timing assumption would thus succeed in giving a bound on the equilibrium payoff set, but only the trivial bound. In contrast, as we will see the bound we investigate here is non-trivial and is in fact tight for some classes of games.

4 Main Results

Throughout the paper, we impose the mild assumption that there is some correlated action profile that all players have strict incentives to approve when disapproval is punished at the correlated minimax level. It will follow from our results that this is a necessary condition for the existence of a repeated game equilibrium where all players receive payoffs strictly above their minimax payoffs.

Formally, let $\underline{u}_i = \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$ be player i's correlated minimax payoff, and let $\left(a_i^{\min}, \alpha_{-i}^{\min}\right)$ be a solution for the problem of minimaxing player i. For $\alpha \in \Delta(A)$, let $d_i(\alpha) = \max_{a_i} u_i\left(a_i, \text{marg}_{-i}\alpha\right)$ be player i's maximum deviation payoff against α. We impose the following:

Assumption There exists $\hat{\alpha} \in \Delta(A)$ such that, for all i, we have

$$u_i(\hat{\alpha}) > (1-\delta)\, d_i(\hat{\alpha}) + \delta \underline{u}_i. \tag{1}$$

This assumption holds whenever G admits a static correlated equilibrium α such that $u_i(\alpha) > \underline{u}_i$ for all i. More generally, it is equivalent to the condition that

$$\delta > \hat{\delta} := \min_{\alpha \in \Delta(A)} \max_{i \in I} \frac{d_i(\alpha) - u_i(\alpha)}{d_i(\alpha) - \underline{u}_i},$$

[2] This logic is similar to, for example, Myerson's (1991; Section 6.1) discussion of games with contracts.

and can thus be viewed as a mild lower bound on the discount factor.

For comparison with the assumptions in our earlier work, note that

$$\hat{\delta} \leq \delta^* := \min_{v \in \mathcal{F}} \max_{i \in I} \frac{d_i}{d_i + v_i - \underline{u}_i},$$

where

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

In Sugaya and Wolitzky (2016), we showed that mediated perfect monitoring is the optimal information structure in two-player repeated games with $\delta > \delta^*$. The condition $\delta > \hat{\delta}$ is thus weaker than the condition $\delta > \delta^*$ from our earlier paper, and indeed the condition $\delta > \hat{\delta}$ is always necessary for the existence of an equilibrium with strictly above-minmax payoffs (while the condition $\delta > \delta^*$ is not). We thus view the condition $\delta > \hat{\delta}$ as a non-triviality assumption, rather than a substantive assumption on the discount factor. Moreover, the results in the current paper make no assumption on the number of players in the game.

4.1 Upper Bound

Our first main result says that the equilibrium payoff set in the repeated delegation game is an upper bound on the equilibrium payoff set in the unmediated repeated game with any private monitoring structure.

Theorem 1 *For any any discount factor δ and any monitoring structure p, the closure of $E_{del} (\delta)$ is an upper bound for $E (\delta, p)$: $\overline{E_{del} (\delta)} \supset E (\delta, p)$.*

Theorem 1 gives an off-the-shelf bound on the Nash equilibrium payoff set for repeated games with imperfect private monitoring.

4.2 Characterization

Our second main result is a tractable characterization of $\overline{E_{del} (\delta)}$. As will be seen, it is without loss to assume that μ is history-independent on the equilibrium path: that is, there exist $\alpha_t \in \Delta (A_t)$ such that $\mu_t(h_m^{t-1}) = \alpha_t$ for every h_m^{t-1} with $c_\tau = (approve_i)_{i \in I}$ for all

$\tau \leq t-1$. Moreover, it is also without loss to assume that if player i unilaterally deviates from the equilibrium (i.e., unilaterally disapproves the mediator's recommended mapping), then the mediator punishes player i at the correlated minimax level. Therefore, a payoff vector $v \in \mathbb{R}^{|I|}$ is supportable in an equilibrium of the delegation game if and only if there exists an action sequence $(\alpha_t)_{t=1}^{\infty}$ for which approval is incentive compatible, given that unilateral disapproval is punished at the correlated minimiax level.

Theorem 2 *For $v \in \mathbb{R}^{|I|}$, $v \in \overline{E_{del}(\delta)}$ if and only if there exists $(\alpha_t)_{t=1}^{\infty}$ such that, for all i and t, we have*

$$(1-\delta)\sum_{\tau \geq t} \delta^{\tau-t} u_i(\alpha_\tau) \geq (1-\delta)\, d_i(\alpha_t) + \delta \underline{u}_i \tag{2}$$

and $v = (1-\delta)\sum_{t\geq 1} \delta^{t-1} u(\alpha_t)$.

Finally, a straightforward application of the methods of APS (modified by setting punishment payoffs at the correlated minimax level) gives the following recursive characterization of $\overline{E_{del}(\delta)}$: Given a set of payoff vectors $W \subset \mathbb{R}^{|I|}$, let $V(W)$ be the set of payoff vectors v such that there exist $\alpha \in \Delta(A)$ and $w \in W$ with

$$(1-\delta)\, u_i(\alpha) + \delta w_i \geq (1-\delta)\, d_i(\alpha) + \delta \underline{u}_i \quad \text{for all } i. \tag{3}$$

Then $\overline{E_{del}(\delta)}$ is the largest finite fixed point of the operator V.

5 Discussion

5.1 Delegation vs. Mediation

The reader may wonder why we need to introduce the repeated delegation game. In particular, an obvious alternative would be to consider a mediated repeated game, as in Forges (1986) and Myerson (1986). In a mediated repeated game, in each period t the mediator recommends a pure action $r_{i,t}$ to player i, rather than a mapping from histories to correlated action profiles. Player i remains free to take any action $a_{i,t}$. The mediator observes (r_t, a_t), while player i observes only her own recommendation and the realized action profile, $(r_{i,t}, a_t)$.

Let $E_{med}(\delta)$ be the sequential equilibrium payoff set in this mediated repeated game with perfect monitoring of actions.

In general, the set $E_{med}(\delta)$ does not give an upper bound for $E(\delta, p)$:

Claim 1 (Sugaya and Wolitzky (2016), Proposition 1) *For some game G and discount factor δ, $\overline{E_{med}(\delta)} \not\supset E(\delta, p)$.*[3]

The intuition for this result is that, compared to perfect monitoring with mediation, imperfect private monitoring has the advantage of "pooling players' information sets." More precisely, under perfect monitoring of actions, a player can perfectly infer what recommendations were made to the other players on the equilibrium path, and can tailor potential deviations to these recommendations. However, under imperfect private monitoring, a player cannot infer her opponents' recommendations—if these recommendations are stochastic along the equilibrium path—and thus has access to coarser information when contemplating a deviation. This difference can make it easier to sustain a given stochastic equilibrium path under private monitoring.

In contrast, in the repeated delegation game, it is without loss to restrict attention to equilibria with deterministic (history-independent) recommendations on the equilibrium path. Imperfect private monitoring thus has no advantage over repeated delegation.

To see why it is without loss to consider history-independent recommendations in the repeated delegation game, imagine that the mediator recommends a mapping $\mu|_{h_m^t}$ after one on-path history h_m^t and recommends $\mu|_{\tilde{h}_m^t}$ after another on-path history $\tilde{h}_m^t$. Suppose the mediator instead recommends the mapping

$$\mu^* := \frac{\Pr\left(h_m^t\right)}{\Pr\left(h_m^t\right) + \Pr\left(\tilde{h}_m^t\right)} \mu|_{h_m^t} + \frac{\Pr\left(\tilde{h}_m^t\right)}{\Pr\left(h_m^t\right) + \Pr\left(\tilde{h}_m^t\right)} \mu|_{\tilde{h}_m^t}$$

after both h_m^t and $\tilde{h}_m^t$, and also employs such a history-independent mapping at all subsequent histories. Then, since it is optimal for player i to approve $\mu|_{h_m^t}$ at history h_m^t and to approve $\mu|_{\tilde{h}_m^t}$ at history $\tilde{h}_m^t$, it is also optimal for her to approve μ^* at history h_m^t or $\tilde{h}_m^t$.

[3] The same result holds when $E(\delta, p)$ is replaced with the *sequential* equilibrium payoff set in game G with discount factor δ and monitoring structure p.

Expected payoffs condition on reaching either h_m^t or $\tilde{h}_m^t$ are also unchanged. The mediator can thus replicate any equilibrium distribution over actions using only history-independent recommendations.

Note that the same argument does not apply with mediation rather than delegation. With mediation, a player must separately "approve" every action in the support of the mediator's on-path recommendation. In order to incentivize the approval of an action for which a player faces a high instantaneous deviation gain, the mediator must promise a high continuation payoff after this action is recommended. The resulting equilibrium path must therefore be stochastic and history-dependent. In contrast, with delegation, approval is required only prior to the mediator's randomization. Continuation payoffs therefore do not need to be tailored to realized actions, and history-independent recommendation strategies are without loss.

Finally, this key advantage of delegation over mediation would also be shared by the repeated game with observable mixed strategies. There are two differences between delegation and observable mixed strategies. First, the delegation game allows for correlated mixtures. This is needed to bound $E(\delta, p)$ because players' actions can be correlated in repeated games with correlated private signals Second, in the delegation game the mediator recommends a *mapping* from history to correlated actions. We explain the role of this second difference next.

5.2 Why Delegation to a Mapping?

A simpler version of the delegation game would have the mediator recommend correlated actions rather than mappings from histories to correlated actions. The advantage of the "mapping" formulation is that players may disagree about the history off-path, and the mapping formulation lets the mediator exploit this disagreement to punish deviators more harshly than would be possible with the "action" formulation.

Specifically, consider the "action" delegation game where, given history h_m^t, the mediator publicly recommends a correlated action $\mu_t(h_m^t)$. Let $E_{act}(\delta)$ be the sequential equilibrium payoff set of this game.

Again, the set $E_{act}(\delta)$ may not give an upper bound for $E(\delta, p)$:

Claim 2 *For some game G and discount factor δ, $\overline{E_{act}(\delta)} \not\supset E(\delta, p)$.*[4]

Proof. Let G be given by the following payoff matrix, and let $\delta = 1/3$:

	A	B	C
A	$1, 1$	$-6, -6$	$0, 1-\eta$
B	$-6, -6$	$1, 1$	$0, 1-\eta$
A'	$2, -6$	$-5, -6$	$0, 1-\eta$
B'	$-5, -6$	$2, -6$	$0, 1-\eta$

We will show that $(1, 1) \in E(p, \delta)$ for some private monitoring structure (Y, p) but $(1, 1) \notin \overline{E_{act}(\delta)}$.

Consider the private monitoring structure given by $Y_1 = Y_2 = \{\alpha, \beta\}$ and

$$\begin{aligned} p(\alpha, \alpha|a_2) &= p(\beta, \beta|a_2) = \frac{1}{2} \text{ if } a_1 = A, B; \\ p(\alpha, \alpha|a_2) &= p(\beta, \beta|a_2) = p(\alpha, \beta|a_2) = p(\beta, \alpha|a_2) = \frac{1}{4} \text{ if } a_1 = A', B'. \end{aligned}$$

Consider the following strategy for each player: In period 1, play A. In subsequent periods, if the previous signal is α, play A; otherwise, play B. The resulting strategy profile yields payoffs $(1, 1)$. We claim that it is a sequential equilibrium strategy profile.

Player 2's incentives are trivial: at every history (on or off path), he believes with probability 1 that player 1's last signal was the same as his own, so he plays a best response.

Player 1's non-trivial incentive constraint is that she must prefer A to A' after signal α, and must prefer B to B' after signal β. If she deviates in some period t, she receives an instantaneous gain of 1, but causes player 2 to mix half-half between A and B in period $t + 1$, indepedently of her own signal. Player 1's expected payoff in period $t + 1$ is therefore $(1/2)(1) + (1/2)(-6) = -5/2$ following a deviation in period t, while this expected payoff is 1 on path. By the one-shot deviation principle, following a deviation in period t player 1 plays A or B in period $t + 1$, and payoffs are again $(1, 1)$ in every period starting in period

[4] Again, the same result holds for sequential equilibrium, as is clear from the proof.

$t+1$. Thus, because

$$\delta(1) = \frac{1}{3} > \frac{1}{6} = 1 - \delta\left(\frac{5}{2}\right),$$

player 1's incentive constraint is satisfied.

By Theorem 1, this implies that $(1,1) \in \overline{E_{med}(\delta)}$. Hence for each $\eta > 0$, Assumption (1) holds.

Now consider repeated "action" delegation. Since player 2's minmax payoff is $1-\eta$ and his maximum feasible payoff is 1, player 2 receives payoff at least $u_2(\eta)$ in every period along the path of play of any sequential equilibrium, where $u(\eta)$ is the solution for

$$(1-\delta)\, u_2(\eta) + \delta \times 1 = 1 - \eta \Leftrightarrow u_2(\eta) = \frac{1-\delta-\eta}{1-\delta}.$$

Fix $\varepsilon > 0$, and suppose toward a contradiction that for each η, there is a sequential equilibrium where player 1's payoff exceeds $1-\varepsilon$. Player 1's greatest payoff consistent with player 2 receiving payoff $u_2(\eta)$ is

$$u_1(\eta) = 1 + \underbrace{\frac{1}{7}}_{\substack{\text{slope of the line segment} \\ \text{connecting (1,1) and (2,−6)}}} (1 - u_2(\eta)) = 1 + \frac{1}{7}\frac{\eta}{1-\delta}.$$

Hence, in such an equilibrium the instantaneous payoff of player 1 in period 1 must exceeds $v_1(\eta)$ such that

$$(1-\delta)\, v_1(\eta) + \delta\left(1 + \frac{1}{7}\frac{\eta}{1-\delta}\right) = 1 - \varepsilon \Leftrightarrow v_1(\eta) = \frac{1}{1-\delta}\left(1 - \varepsilon - \delta\left(1 + \frac{1}{7}\frac{\eta}{1-\delta}\right)\right).$$

Since player 2 receives payoff at least $u_2(\eta)$, probability that (A,A) or (B,B) is played in period 1 must exceeds $p(\eta)$ with

$$\begin{cases} p(\eta) + q(\eta)\, 2 = v_1(\eta) \\ (1 - q(\eta)) + q(\eta)(-6) = u_2(\eta) \end{cases}$$

$$\Rightarrow \quad p(\eta) = \frac{1}{1-\delta}\left(1 - \varepsilon - \delta\left(1 + \frac{1}{7}\frac{\eta}{1-\delta}\right)\right) - \frac{2}{7}\frac{\eta}{1-\delta},$$

where $q(\eta)$ is the probability of (A', A) or (B', B). Hence player 1's instantaneous gain from deviating from A to A' or from B to B' in period 1 is at least $p(\eta)$.

Now, consider the harshest punishment that the mediator can conduct after player 1's deviation. For each h_m^t, suppose the mediator asks for the approval for α. Since player 2's instantaneous payoff is bounded from below by $u_2(\eta)$, player 1's payoff given α should be no less than[5]

$$w_1(\eta) = 0 - \underbrace{\frac{6}{6-(1-\eta)}}_{\substack{\text{slope of the line segment} \\ \text{connecting } (0,1-\eta) \text{ and } (-6,-6)}} (1-\eta-u_2(\eta)) = -6\delta\frac{\eta}{(1-\delta)(\eta+5)}.$$

This implies that player 1's expected continuation payoff after she deviates in period 1 is at least $w_1(\eta)$. But, for ε and η sufficiently small,

$$\delta(1) < (1-\delta)p(\eta) + \delta w_1(\eta),$$

as $\delta < 1/2$, so player 1 prefers to deviate in period 1. This gives the desired contradiction. ■

Intuitively, with private monitoring, it can be possible to induce one player to punish another without realizing that a deviation has occured. This allows for punishments where the punisher herself receives less than her minmax payoff. With action delegation and observable actions, this may no longer be possible.[6] In contrast, with delegation to the mapping, if players believe that they are on equilibrium path, then they are willing to approve the recommended mapping even if it specifies costly punishment in the event that a deviation has in fact occured.

It can however be shown that $E_{del}(\delta) = E_{act}(\delta)$ whenever $\delta > \delta^*$. This follows because it is possible to minimax deviators with mediation when $\delta > \delta^*$, as shown in Sugaya and Wolitzky (2016).

[5] Note that players 1 and 2 may have different beliefs about the continuation play from period $t+1$ on. However, in the action delegation, α becomes common knowledge.

[6] This effect has previously been noted by Kandori (1991), Sekiguchi (2002), and Sugaya and Wolitzky (2016), among others.

5.3 Tightness of the Bound?

It is obvious that, for some games and some monitoring structures, $E(p, \delta)$ is a strict subset of $E_{del}(\delta)$. A more interesting question is whether $E_{del}(\delta)$ a tight bound on $E(p, \delta)$ from the perspective of an observer who does not know the monitoring structure p: that is, for each $v \in E_{del}(\delta)$, does there exist a monitoring structure p such that $v \in E(\delta, p)$? Unfortunately, in general the answer to this question is no.

The logic is similar to the reason why allowing observable mixed strategies can expand the equilibrium set. For example, consider the following game:

	L	R
T	$1, 1$	$0, 2$
B	$1, -1$	$0, 0$

If $\delta < 1/3$, then player 2 cannot play L in the repeated game since the deviation gain of $1-\delta$ would exceed the greatest possible change in continuation payoff (from $\delta(2)$ to $\delta(0)$, noting that 0 is the minmax payoff). Hence, the best payoff for player 1 is 0.

However, with delegation to the mediator, the following is an equilibrium whenever $\delta \geq 1/4$: play $\frac{1}{2}(T, L) + \frac{1}{2}(T, R)$ on the equilibrium path, and play (B, R) after any deviation. This follows because player 2's deviation gain is $(1-\delta)/2$ and his lost continuation payoff from a deviation is $\delta(3/2)$.

Hence, for any $\delta \in (1/4, 1/3)$, player 1's best equilibrium payoff with any monitoring structure is 0, while her best equilibrium payoff with delegation to the mediator is strictly positive.

While $E_{del}(\delta)$ does not give a tight bound in general, it does give a tight bound for *concave games* as defined in Sugaya and Wolitzky (2016b). A concave game is one in which, for each player i,

1. A_i is a compact and convex subset of $\mathbb{R}^n$,
2. $u_i(a)$ is continuous and jointly concave in a, and
3. $d_i(\alpha)$ is jointly convex in α.

In a concave game, replacing a distribution over actions with its expectation increases players' payoffs and reduces their incentives to deviate. Sugaya and Wolitzky (2016b) show that, if under perfect monitoring it is possible to hold a deviator to her minimax payoff, then every $v \in E(\delta, p)$ is Pareto dominated by a subgame perfect equilibrium payoff vector under perfect monitoring, for every private monitoring structure p. The same argument applies to every $v \in E_{del}(\delta)$. Thus, at least as far as efficient payoffs are concerned, in concave games $E_{del}(\delta)$ provides a tight bound on the equilibrium payoff set, which coincides with the equilibrium payoff set in the perfect monitoring repeated game.

5.4 Efficient Equilibria in Symmetric Games: A Simpler Characterization

For symmetric games, a simpler characterization of the best symmetric payoff vector in $E_{del}(\delta)$ is available. In particular, in computing this payoff vector, attention can be restricted to stationary equilibria.[7]

Proposition 1 *If $\delta > \hat{\delta}$, G is symmetric, v is symmetric, and $v \in \overline{E_{del}(\delta)}$, then there exists a symmetric correlated action $\alpha \in \Delta(A)$ such that $u(\alpha) \geq v$ and, for all i,*

$$u_i(\alpha) \geq (1-\delta) d_i(\alpha) + \delta \underline{u}_i.$$

Proof. By Theorem 2, there exists $(\alpha_t)_{t=1}^{\infty}$ such that (2) holds and $v = (1-\delta)\sum_{t \geq 1} \delta^{t-1} u(\alpha_t)$. Let $w(\alpha_t) = \sum_{i \in I} u_i(\alpha_t)$, and let $w = \sup_t w(\alpha_t)$. Note that $(w/|I|, \ldots, w/|I|) \geq v$ and (by summing (2) across players) $w \geq \sum_i [(1-\delta) d_i(\alpha_t) + \delta \underline{u}_i]$ for all t. As A is finite and payoffs are continuous in mixing probabilities, there exists α such that $\sum_{i \in I} u(\alpha) = w$ and $w \geq \sum_i [(1-\delta) d_i(\alpha) + \delta \underline{u}_i]$. Now, let Π be the set of all permutations π on I; for each $\pi \in \Pi$ let α^π be the correlated action given by $\alpha^\pi\left(a_1, \ldots, a_{|I|}\right) = \alpha\left(a_{\pi(1)}, \ldots, a_{\pi(|I|)}\right)$; and let $\bar{\alpha} = (1/|I|!)\sum_{\pi \in \Pi} \alpha^\pi$. Then, for all i, $u_i(\bar{\alpha}) = w/|I|$ and $d_i(\bar{\alpha}) = (1/|I|)\sum_i d_i(\alpha)$. As $w \geq \sum_i [(1-\delta) d_i(\alpha) + \delta \underline{u}_i]$ and $\underline{u}_i$ is the same for all i, it follows that $u_i(\bar{\alpha}) \geq (1-\delta) d_i(\bar{\alpha}) + \delta \underline{u}_i$ for all i. ■

[7] A related result for concave games appears in Sugaya and Wolitzky (2016b).

5.5 The Characterization in an Example

We illustrate Theorem 2 and Proposition 1 with an application to a simple repeated public good provision game.

Consider the symmetric game with $A_i = \mathbb{R}_+$ and

$$u_i(a) = f\left(\sum_{j=1}^{|I|} a_j\right) - a_i,$$

where f is an increasing and concave function with $f(0) = 0$ and $f'(x) < 1$ for all x.[8] Note that $a_i = 0$ is dominant in the stage game. Note also that this is not a concave game in the sense of the previous section, because the deviation payoff $d_i(\alpha)$ is not jointly convex in actions: since f is concave, the deviation payoff can be larger if $\sum_{j \neq i} a_j$ is known to be intermediate than if it could be either small or large.

We wish to characterize the best symmetric payoff vector in $\overline{E_{del}(\delta)}$. By Proposition 1, we know that this is given by a symmetric stationary equilibrium α, with deviators punished at the minimax payoff of 0.

We first show that α only puts weight on vectors a with $a_i = 0$ for all but a single player i. That is, it is never optimal to have two players work at the same time.

To see this, let $a_{-i} = \sum_{j \neq i} a_j$, and note that player i's deviation gain from α is given by

$$\mathbb{E}[a_i] + \mathbb{E}[f(a_{-i})] - \mathbb{E}[f(a_i + a_{-i})].$$

A necessary condition for α to be optimal is thus that it maximizes $\mathbb{E}[f(a_i + a_{-i}) - f(a_{-i})]$ for given $\mathbb{E}[a_i]$ (for each i) and $\mathbb{E}[f(a_i + a_{-i})]$. As f is concave, this is given by having player i work only when no one else is working.

It remains only to find the optimal effort level a^* such that randomly having one player work at level a^* each period is optimal. (Of course, a player does not know if she will be the one to work when she approves the recommendation.) Assuming a^* is below the first-best

[8] This game has a continuum of actions. It will become clear that discretizing and bounding A_i would have no effect on the analysis.

level (defined by $f'(na) = 1/n$), this is given by

$$f(a^*) - \frac{1}{n}a^* = (1-\delta)\frac{n-1}{n}f(a^*),$$

or

$$a^* = f(a^*) + \delta(n-1)f(a^*).$$

The best symmetric payoff vector in $\overline{E_{del}(\delta)}$ thus gives each player payoff

$$f(a^*) - \frac{1}{n}a^* = (n-1)\left(\frac{1}{n}+\delta\right)f(a^*).$$

This payoff is therefore an upper bound on the best symmetric payoff attainable in a repeated game Nash equilibrium for any private monitoring structure.

While this payoff bound is permissive, it is not obvious (to us, at least) how to give a tighter bound in this game. One obvious upper bound on each player's action is given by the value $\bar{a}$ that solves

$$(1-\delta)[\bar{a} - f(\bar{a})] = \delta f((n-1)\bar{a}),$$

or

$$\bar{a} = f(\bar{a}) + \frac{\delta}{1-\delta}f((n-1)\bar{a}).$$

This bound follows because a player's minmax payoff is 0, the instantaneous gain from deviating from $a_i = \bar{a}$ to $a_i = 0$ is at least $(1-\delta)[\bar{a} - f(\bar{a})]$, and—if $\bar{a}$ is the least upper bound on a player's equilibrium action—a player's continuation payoff is at most $\delta f((n-1)\bar{a})$. Comparing a^* to $\bar{a}$, we see that $a^* < \bar{a}$ if and only if δ is above some cutoff $\bar{\delta}$ (noting that $(n-1)f(a) > f((n-1)a)$ for all $a > 0$, as f is concave and $f(0) = 0$). Thus, for sufficiently patient players the payoff bound coming from our theorems is tighter than the obvious bound.

6 Proof of Theorems

6.1 A Useful Lemma

Lemma 1 *There exists $\bar{\varepsilon} > 0$ such that, for each $\varepsilon < \bar{\varepsilon}$, it is an obedient sequential equilibrium strategy profile in the delegation game for the players to always approve the mediator's recommendation and for the mediator to recommended the following mapping in period t:*

- *If all players have always approved prior to time t, then $\mu\left(h_m^t\right) = (1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a \in A} a$.*
- *If the set of players who have ever disapproved prior to time t is the singleton $\{i\}$ for some $i \in I$, then $\mu\left(h_m^t\right) = \left(a_i^{\min}, \alpha_{-i}^{\min}\right)$.*
- *If the set of players who have ever disapproved prior to time t is $J \subset I$ with $|J| \geq 2$, then $\mu\left(h_m^t\right) = \left(\left(a_j^{\min}\right)_{j \in J}, \alpha_{-J}^{\text{mix}}\right)$, where α_{-J}^{mix} is the correlated action that assigns probability $1/\left|A_{-J}\right|$ to each action profile $a_{-J} \in A_{-J}$.*

Proof. Since the on-path correlated action $(1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a \in A} a$ has full support, a player who has always previously approved believes that, (i) if she approves then $(1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a \in A} a$ will be played forever; and (ii) if she disapproves and plays a_i then

$$\left(a_i, \operatorname{marg}_{-i}\left((1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a \in A} a\right)\right)$$

will be played in the current period and $\left(a_i^{\min}, \alpha_{-i}^{\min}\right)$ will be played in every future period. For sufficiently small ε, (1) implies that approval is optimal

If a player has ever disapproved, she believes that her opponents will play $\alpha_{-i}^{\min}$ in every subsequent period regardless of her own behavior, and she believes that she will play $a_i^{\min}$ in every subsequent period if she approves. Since $a_i^{\min}$ is a best response to $\alpha_{-i}^{\min}$, approval is optimal.

Thus, by one-shot deviation principle, it is optimal to always approve the mediator's recommendation. ■

Henceforth, we fix $\bar{\varepsilon}$ such that the conclusion of the lemma applies.

6.2 Proof of Theorem 2

The proof of Theorem 2 consists of the following two propositions:

Given a strategy for the mediator μ, let $\mathbb{E}^{\mu}[\cdot]$ denote expectation with respect to the induced distribution over action profiles.

Proposition 2 *Let μ be an obedient sequential equilibrium strategy for the mediator in the delegation game. Then, letting $\alpha_t = \mathbb{E}^{\mu}[\mu(h_m^t)]$ for all t, (2) holds for all i and t.*

Proof. Since a player's continuation payoff cannot fall below her minimax payoff, approval is optimal at on-path history h_i^t only if

$$(1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}\mathbb{E}^{\mu}\left[u_i\left(\mu\left(h_m^{\tau}\right)\right)|h_i^t\right] \geq (1-\delta)\max_{a_i}\mathbb{E}^{\mu}\left[u_i\left(a_i, \mathrm{marg}_{-i}\mu\left(h_m^t\right)\right)|h_i^t\right] + \delta\underline{u}_i. \quad (4)$$

By the law of iterated expectations,

$$\mathbb{E}^{\mu}\left[(1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}\mathbb{E}^{\mu}\left[u_i\left(\mu\left(h_m^{\tau}\right)\right)|h_i^t\right]\right] = (1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}\mathbb{E}^{\mu}\left[u_i\left(\mu\left(h_m^{\tau}\right)\right)\right].$$

By convexity of the max operator and Jensen's inequality,

$$\begin{aligned}\mathbb{E}^{\mu}\left[\max_{a_i}\mathbb{E}^{\mu}\left[u_i\left(a_i, \mathrm{marg}_{-i}\mu\left(h_m^t\right)\right)|h_i^t\right]\right] &\geq \max_{a_i}\mathbb{E}^{\mu}\left[\mathbb{E}^{\mu}\left[u_i\left(a_i, \mathrm{marg}_{-i}\mu\left(h_m^t\right)\right)|h_i^t\right]\right] \\ &= \max_{a_i}\mathbb{E}^{\mu}\left[u_i\left(a_i, \mathrm{marg}_{-i}\alpha_{,t}\right)\right] = d_i\left(\alpha_t\right).\end{aligned}$$

Hence, taking the expectation of both sides of (4) with respect to h_i^t yields (2). ■

The following proposition proves the converse:

Proposition 3 *If $(\alpha_t)_{t=1}^{\infty}$ satisfies (2) for all i and t, then $(1-\delta)\sum_{t\geq 1}\delta^{t-1}u_i(\alpha_t) \in \overline{E_{del}(\delta)}$.*

Proof. It suffices to show that, for any $\eta > 0$, there exists an incentive-compatible strategy for the mediator μ such that $\max_i\left|v_i - (1-\delta)\sum_{t\geq 1}\delta^{t-1}\mathbb{E}^{\mu}\left[u_i(a_t)\right]\right| < \eta$.

Fix $\varepsilon \in (0, \bar{\varepsilon})$, and consider the following strategy for the mediator. For every t, the mediator recommends the mapping $\mu_t : H_m^t \to \Delta(A)$ defined as follows:

- If all players have always approved prior to time t, then $\mu_t\left(h_m^t\right) = (1-\varepsilon)\,\alpha_t + \varepsilon\left[(1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a\in A} a\right]$.

- If the set of players who have ever disapproved prior to time t is the singleton $\{i\}$ for some $i \in I$, then $\mu_t\left(h_m^t\right) = \left(a_i^{\min}, \alpha_{-i}^{\min}\right)$.

- If the set of players who have ever disapproved prior to time t is $J \subset I$ with $|J| \geq 2$, the $\mu_t\left(h_m^t\right) = \left(\left(a_j^{\min}\right)_{j\in J}, \alpha_{-J}^{\text{mix}}\right)$.

Since the on-path correlated action has full support, a player who has always approved believes that all players have always approved, and a player who has ever disapproved believes that she is the only player who has ever disapproved.

If a a player has ever disapproved, it is optimal for her to approve in the future by the same argument as in Lemma 1. We are thus left to consider her a player's incentives when she has always approved.

In this case, for every i and t,

$$
\begin{aligned}
& (1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}\mathbb{E}^{\mu}\left[u_i\left(a_\tau\right)|h_i^t\right] \\
= \;& (1-\varepsilon)(1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}u_i\left(\alpha_t\right) + \varepsilon(1-\delta)\sum_{\tau\geq t}\delta^{\tau-t}u_i\left((1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a\in A} a\right) \\
\geq \;& (1-\varepsilon)\left[(1-\delta)\,d_i\left(\alpha_t\right) + \delta\underline{u}_i\right] + \varepsilon\left[(1-\delta)\,d_i\left((1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a\in A} a\right) + \delta\underline{u}_i\right] \\
\geq \;& (1-\delta)\,d_i\left((1-\varepsilon)\,\alpha_t + \varepsilon\left[(1-\varepsilon)\,\hat{\alpha} + \frac{\varepsilon}{|A|}\sum_{a\in A} a\right]\right) + \delta\underline{u}_i,
\end{aligned}
$$

where the first inequality follows by (2) and Lemma 1 (given the assumption that $\varepsilon < \bar{\varepsilon}$), and the second inequality follows because the function $d_i(\cdot)$ is convex (as the max operator is convex). Hence, it is optimal to approve the mediator's recommendation for all $\varepsilon \in (0, \bar{\varepsilon})$.

Finally, for every $\eta > 0$ there exists $\varepsilon \in (0, \bar{\varepsilon})$ such that $\max_i \left|v_i - (1-\delta)\sum_{t\geq 1}\delta^{t-1}\mathbb{E}^{\mu}\left[u_i\left(a_t\right)\right]\right| < \eta$. ■

6.3 Proof of Theorem 1

Fix a Nash equilibrium σ of the repeated game with some monitoring structure, and let $\alpha_t = \mathbb{E}^{\sigma}[a_t]$ be the induced distribution over period t actions. By Proposition 3, it suffices to show that $(\alpha_t)_{t=1}^{\infty}$ satisfies (2) for all i and t.

Since a player's continuation payoff cannot fall below her minimax payoff, the fact that σ is a Nash equilibrium implies that, for every on-path history h_i^t,

$$(1-\delta)\sum_{\tau \geq t} \delta^{\tau - t} \mathbb{E}^{\sigma}\left[u_i(a_t)\,|h_i^t\right] \geq (1-\delta)\max_{a_i} \mathbb{E}^{\sigma}\left[u_i(a_i, a_{-i})\,|h_i^t\right] + \delta \underline{u}_i.$$

As in the proof of Proposition 2, taking the expectation of both sides of this inequality with respect to h_i^t, collapsing the iterated expectation on the left-hand side, and using the convexity of the max operator and Jensen's inequality to move the outer expectation inside the max operator on the right-hand side yields (2).