Theoretical Economics 12 (2017), 691–729 1555-7561/20170691

Bounding equilibrium payoffs in repeated games with private monitoring

Takuo Sugaya
Graduate School of Business, Stanford University

Alexander Wolitzky
Department of Economics, MIT

We provide a simple sufficient condition for the existence of a recursive upper bound on (the Pareto frontier of) the sequential equilibrium payoff set at a fixed discount factor in two-player repeated games with imperfect private monitoring. The bounding set is the sequential equilibrium payoff set with perfect monitoring and a mediator. We show that this bounding set admits a simple recursive characterization, which nonetheless necessarily involves the use of private strategies. Under our condition, this set describes precisely those payoff vectors that arise in equilibrium for some private monitoring structure if either nonstationary monitoring or communication is allowed.

Keywords. Repeated games, private monitoring.

JEL classification. C72, C73.

1. Introduction

Like many dynamic economic models, repeated games are typically studied using recursive methods. In an incisive paper, Abreu et al. (1990; henceforth APS) recursively characterized the perfect public equilibrium payoff set at a fixed discount factor in repeated games with imperfect public monitoring. Their results (along with related contributions by Fudenberg et al. (1994) and others) led to fresh perspectives on problems like collusion (Green and Porter 1984, Athey and Bagwell 2001), relational contracting (Levin 2003), and government credibility (Phelan and Stacchetti 2001). However, other important environments—like collusion with secret price cuts (Stigler 1964) or relational contracting with subjective performance evaluations (Levin 2003, MacLeod 2003, Fuchs 2007)—involve imperfect *private* monitoring, and it is well known that the methods of

Takuo Sugaya: tsugaya@stanford.edu
Alexander Wolitzky: wolitzky@mit.edu
This replaces our earlier paper, "Perfect Versus Imperfect Monitoring in Repeated Games." For helpful comments, we thank Dilip Abreu, Gabriel Carroll, Henrique de Oliveira, Glenn Ellison, Drew Fudenberg, Michihiro Kandori, George Mailath, Hitoshi Matsushima, Stephen Morris, Ilya Segal, Tadashi Sekiguchi, Andrzej Skrzypacz, Satoru Takahashi, Juuso Toikka, Thomas Wiseman, the anonymous referees, and participants in several seminars.

DOI: 10.3982/TE2270

APS do not easily extend to such settings (Kandori 2002). Whether the equilibrium payoff set in repeated games with private monitoring exhibits any tractable recursive structure at all is thus a major question.

In this paper, we do *not* make any progress toward giving a recursive characterization of the sequential equilibrium payoff set in a repeated game with a *given* private monitoring structure. Instead, working in the context of two-player games, we provide a simple condition for the existence of a recursive upper bound on the Pareto frontier of this set.[1] The key feature of our bound is that it is tight from the perspective of an observer who does not know the monitoring structure under which the game is being played: that is, our bound *characterizes* the set of payoffs that can arise in equilibrium for *some* monitoring structure. In other words, from the perspective of an observer who knows the monitoring structure, our results give an upper bound on how well the players can do; while from the perspective of an observer who does not know the monitoring structure, our results exactly characterize how well they can do. Which of these two perspectives on our results is more relevant for a particular application thus depends on the observability of the monitoring structure to an outsider, which can be expected to vary from application to application.

The set we use to upper-bound the equilibrium payoff set with private monitoring is the equilibrium payoff set with perfect monitoring and a mediator (*mediated perfect monitoring*). We do not take a position on the realism of allowing a mediator, and instead view the model with a mediator as a purely technical device that is useful for bounding equilibrium payoffs with private monitoring. We thus show that the equilibrium payoff set with private monitoring admits a recursive upper bound by establishing two main results:

(i) Under a simple condition, the equilibrium payoff set with mediated perfect monitoring is an upper bound on the equilibrium payoff set with any private monitoring structure.

(ii) The equilibrium payoff set with mediated perfect monitoring has a recursive structure.

It might seem surprising that any conditions at all are needed for the first of these results, as one might think that improving the precision of the monitoring structure and adding a mediator can only expand the equilibrium set. But this is not the case: giving a player more information about her opponents' past actions splits her information sets and thus gives her new ways to cheat, and indeed we show by example that (unmediated) imperfect private monitoring can sometimes outperform (mediated) perfect monitoring. In other words, perfect monitoring is not necessarily the optimal monitoring structure in a repeated game, even if it is advantaged by giving the players access to a mediator.

Our sufficient condition for mediated perfect monitoring to outperform any private monitoring structure is that there is a feasible payoff vector v such that no player i is

[1]For conciseness, henceforth we will say that a set of payoff vectors X *upper-bounds* set Y if every payoff vector $y \in Y$ is Pareto dominated by some payoff vector $x \in X$. This is obviously not the same as $X \supseteq Y$.

tempted to deviate if she gets continuation payoff v_i when she conforms and is minmaxed when she deviates. This is a joint restriction on the stage game and the discount factor, and it is essentially always satisfied when players are at least moderately patient. (However, they need not be patient enough for the folk theorem to apply.) Under this condition, we show that the Pareto frontier of the equilibrium payoff set under mediated perfect monitoring coincides with that under the *universal monitoring structure* that arises when the mediator perfectly observes all actions but each player observes only her own actions. Our first main result follows because, as its name suggests, the universal monitoring structure embeds any private monitoring structure.

To understand our second main result, recall that, in repeated games with perfect monitoring *without* a mediator, all strategies are public, so the sequential (equivalently, subgame perfect) equilibrium set coincides with the perfect public equilibrium set, which was recursively characterized by APS. However, *with* a mediator—who makes private action recommendations to the players—private strategies play a crucial role, and APS's characterization does not apply. Nonetheless, under the sufficient condition for our first result, a recursive characterization is obtained by replacing APS's generating operator B with what we call the *minmax-threat generating operator* $\tilde{B}$: for any set of continuation payoffs W, the set $\tilde{B}(W)$ is the set of payoffs that can be attained when on-path continuation payoffs are drawn from W and deviators are minmaxed. To see why deviators can always be minmaxed in the presence of a mediator—and also why private strategies cannot be ignored—suppose that the mediator recommends a target action profile $a \in A$ with probability $1 - \varepsilon$, while recommending every other action profile with probability $\varepsilon/(|A| - 1)$. Suppose further that if some player i deviates from her recommendation, the mediator then recommends that her opponents minmax her in every future period. In such a construction, player i's opponents never learn that a deviation has occurred, and they are therefore always willing to follow the recommendation of minmaxing player i.[2] (This construction clearly relies on private strategies: if the mediator's recommendations were public, players would always see when a deviation occurs, and they then might not be willing to minmax the deviator.)

We consider several extensions of our results. Perhaps most importantly, we establish two senses in which the equilibrium payoff set with mediated perfect monitoring is a *tight* upper bound on the equilibrium payoff set, from the perspective of an observer who does not know the monitoring structure. First, mediated perfect monitoring with a given strategy of the mediator's itself induces a *nonstationary* monitoring structure, meaning that the distribution of signals can depend on everything that has happened in the past, rather than only on current actions. Thus, our upper bound is trivially tight from the perspective of an observer who finds nonstationary monitoring structures possible. Second, restricting attention to standard, *stationary* monitoring structures—where the signal distribution depends only on the current actions—we show that the mediator can be dispensed with if the players have access to an ex ante correlating device and cheap talk communication. Hence, our upper bound is also tight from the

[2]In this construction, the mediator *virtually implements* the target action profile. For other applications of virtual implementation in games with a mediator, see Lehrer (1992), Mertens et al. (2015, IV.4.b), Renault and Tomala (2004), Rahman and Obara (2010), and Rahman (2012).

perspective of an observer who finds only stationary monitoring structures possible, if she also accepts the possibility of ex ante correlation and cheap talk.

This paper is not the first to develop recursive methods for repeated games with imperfect private monitoring. Kandori and Matsushima (1998) augment private monitoring repeated games with opportunities for public communication among the players and provide a recursive characterization of a subset of equilibrium payoffs that is large enough to yield a folk theorem. Tomala (2009) gives related results when the repeated game is augmented with a mediator rather than only public communication. Neither paper provides a recursive upper bound on the entire sequential equilibrium payoff set for a fixed discount factor.[3] Amarante (2003) does give a recursive characterization of the equilibrium payoff set in private monitoring repeated games, but the state space in his characterization is the set of repeated game histories, which grows over time. Phelan and Skrzypacz (2012) and Kandori and Obara (2010) develop recursive methods for *checking* whether a given finite-state strategy profile is an equilibrium in a private monitoring repeated game.

Awaya and Krishna (2015) and Pai et al. (2014) derive bounds on payoffs in private monitoring repeated games as a function of the monitoring structure. The bounds in these papers come from the observation that if an individual's actions can have only a small impact on the distribution of signals, then the shadow of the future can have only a small effect on her incentives. In contrast, our payoff bounds apply for all monitoring structures, including those in which individual actions have a large impact on the signal distribution.

Finally, we have emphasized that our results can be interpreted either as giving an upper bound on the equilibrium payoff set in a repeated game for a *particular* private monitoring structure or as characterizing the set of payoffs that can arise in equilibrium for *some* private monitoring structure. With the latter interpretation, our paper shares a motivation with Bergemann and Morris (2013), who characterize the set of payoffs that can arise in equilibrium in a static incomplete information game for some information structure. Yet another interpretation of our results is that they establish that information is valuable in mediated repeated games, in that—under our sufficient condition—players cannot benefit from imperfections in the monitoring technology. This interpretation connects our paper to the literature on the value of information in static incomplete information games (e.g., Gossner 2000, Lehrer et al. 2010, Bergemann and Morris 2013).

The rest of the paper is organized as follows. Section 2 describes our models of repeated games with private and mediated perfect monitoring, which are standard. Section 3 gives an example showing that private monitoring can sometimes outperform mediated perfect monitoring. Section 4 develops preliminary results about repeated

[3]Ben-Porath and Kahneman (1996) and Compte (1998) also prove folk theorems for private monitoring repeated games with communication, but they do not emphasize recursive methods away from the $\delta \to 1$ limit. Lehrer (1992), Mertens et al. (2015, IV.4.b), and Renault and Tomala (2004) characterize the communication equilibrium payoff set in *undiscounted* repeated games. These papers study how imperfect monitoring can limit the equilibrium payoff set without discounting, while our focus is on how discounting can limit the equilibrium payoff set independently of the monitoring structure.

games with mediated perfect monitoring. Section 5 presents our first main result: a sufficient condition for mediated perfect monitoring to outperform private monitoring. The proof of this result is deferred to Section 8. Section 6 presents our second main result: a recursive characterization of the equilibrium payoff set with mediated perfect monitoring. Section 7 illustrates the calculation of the upper bound with an example. Section 9 discusses the tightness of our upper bound, as well as partial versions of our results that apply when our sufficient conditions do not hold, as in the case of more than two players. Section 10 concludes. Additional material is available in a supplementary file on the journal website, http://econtheory.org/supp/2270/supplement.pdf.

2. Repeated games with private and mediated perfect monitoring

A finite stage game $G = (I, (A_i, u_i)_{i\in I})$ 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 δ.

2.1 *Private monitoring*

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

A period t history for player i is an element of $H_i^t = (A_i \times Z_i)^{t-1}$, with typical element $h_i^t = (a_{i,\tau}, z_{i,\tau})_{\tau=1}^{t-1}$, where H_i^1 consists of the null history $\varnothing$. A (behavior) strategy for player i is a map $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$. A belief system for player i is a map $\beta_i : \bigcup_{t=1}^{\infty} H_i^t \to \bigcup_{t=1}^{\infty} \Delta(H^t)$ satisfying $\operatorname{supp} \beta_i(h_i^t) \subseteq \{h_i^t\} \times H_{-i}^t$ for all t; we also write $\beta_i(h^t|h_i^t)$ for the probability of h^t under $\beta_i(h_i^t)$. Let $H^t = \prod_{i\in I} H_i^t$.

The solution concept is sequential equilibrium.

Definition 1. An assessment (σ, β) constitutes a *sequential equilibrium* if the following two conditions are satisfied:

(i) Sequential rationality. For each player i and history h_i^t, σ_i maximizes player i's expected continuation payoff at history h_i^t under belief $\beta_i(h_i^t)$.

(ii) Consistency. There exists a sequence of completely mixed strategy profiles (σ^n) such that the following two conditions hold:

(a) The sequence (σ^n) converges to σ (pointwise in t). For all $\varepsilon > 0$ and t, there exists N such that, for all $n > N$, $|\sigma_i^n(h_i^t) - \sigma_i(h_i^t)| < \varepsilon$ for all $i \in I$, $h_i^t \in H_i^t$.

(b) Conditional probabilities converge to β (pointwise in t). For all $\varepsilon > 0$ and t, there exists N such that, for all $n > N$,

$$\left| \frac{\Pr^{\sigma^n}(h_i^t, h_{-i}^t)}{\sum_{\tilde{h}_{-i}^t} \Pr^{\sigma^n}(h_i^t, \tilde{h}_{-i}^t)} - \beta_i(h_i^t, h_{-i}^t \mid h_i^t) \right| < \varepsilon \quad \text{for all } i \in I, h_i^t \in H_i^t, h_{-i}^t \in H_{-i}^t.$$

This relatively permissive definition of consistency (requiring that strategies and beliefs converge only pointwise in t) gives a weakly larger set of equilibrium payoffs to be bounded but also allows more freedom in constructing the bounding equilibria. However, by replacing infinite punishments with long finite punishments, our equilibrium constructions can be modified to satisfy consistency under uniform convergence.

2.2 *Mediated perfect monitoring*

In each period t, the game proceeds as follows: A mediator sends a private message $m_{i,t} \in M_i$ to each player i, where M_i is a finite message set for player i. Each player i takes an action $a_{i,t} \in A_i$. All players and the mediator observe the action profile $a_t \in A$.

A period t history for the mediator is an element of $H_m^t = (M \times A)^{t-1}$, with typical element $h_m^t = (m_\tau, a_\tau)_{\tau=1}^{t-1}$, where H_m^1 consists of the null history. A strategy for the mediator is a map $\mu : \bigcup_{t=1}^{\infty} H_m^t \to \Delta(M)$. A period t history for player i is an element of $H_i^t = (M_i \times A)^{t-1} \times M_i$, with typical element $h_i^t = ((m_{i,\tau}, a_\tau)_{\tau=1}^{t-1}, m_{i,t})$, where $H_i^1 = M_i$.[4] A strategy for player i is a map $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$.

The definition of sequential equilibrium is the same as with private monitoring, except that sequential rationality is imposed (and beliefs are defined) 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.[5] Note that with this definition, an assessment (including the mediator's strategy) (μ, σ, β) is a sequential equilibrium with mediated perfect monitoring if and only if (σ, β) is a sequential equilibrium with the "nonstationary" private monitoring structure where $Z_i = M_i \times A$ and $p_t(\cdot | h_m^{t+1})$ coincides with perfect monitoring of actions with messages given by $\mu(h_m^{t+1})$ (see Section 9.1).

As in Forges (1986) and Myerson (1986), any equilibrium distribution over action paths arises in an equilibrium of the following form:

(i) Messages are action recommendations: $M = A$.

(ii) Obedience/incentive compatibility. At history $h_i^t = ((m_{i,\tau}, a_\tau)_{\tau=1}^{t-1}, m_{i,t})$, player i plays $a_{i,t} = m_{i,t}$.

Without loss of generality, we restrict attention to such *obedient* equilibria throughout.[6] We denote the mediator's action recommendation by $r \in A$.

Finally, we say that a sequential equilibrium with mediated perfect monitoring is *on-path strict* if following the mediator's recommendation is strictly optimal for each player i at every on-path history h_i^t. Let $E_{\text{med}}(\delta)$ denote the set of on-path strict sequential equilibrium payoffs. For the rest of the paper, we slightly abuse terminology by omitting the qualifier "on-path" when discussing such equilibria.

[4] We also occasionally write h_i^t for $(m_{i,\tau}, a_\tau)_{\tau=1}^{t-1}$, omitting the period t message $m_{i,t}$.

[5] The assumption that the mediator cannot tremble does not matter for our results.

[6] Dhillon and Mertens (1996) show that the revelation principle fails for trembling-hand perfect equilibria. Nonetheless, with our machine interpretation of the mediator, the revelation principle applies for sequential equilibrium by precisely the argument of Forges (1986).

3. An illustrative (counter)example

The goal of this paper is to provide sufficient conditions for the equilibrium payoff set with mediated perfect monitoring to upper-bound the equilibrium payoff set with private monitoring. We first provide an illustrative example showing why, in the absence of our sufficient conditions, private monitoring (without a mediator) can outperform mediated perfect monitoring. Readers eager to get to the results can skip this section without loss of continuity.

Consider the repetition of the following stage game, with $\delta = \frac{1}{6}$:

	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	$0, -3$	$0, 3$	$0, 3$

Example 1

Proposition 1. *In Example 1, there is no sequential equilibrium where the players' per-period payoffs sum to more than* 3 *with mediated perfect monitoring, while there is such a sequential equilibrium with some private monitoring structure.*

Let us sketch the proof of Proposition 1. Note that (U, L) is the only action profile where payoffs sum to more than 3. Because δ is low, player 1 (row player, "she") can be induced to play U in response to L only if action profile (U, L) is *immediately* followed by (T, M) with high enough probability: specifically, this probability must exceed $\frac{3}{5}$. With perfect monitoring, player 2 (column player, "he") must then "see (T, M) coming" with probability at least $\frac{3}{5}$ following (U, L), and this probability is so high that player 2 will deviate from M to L (regardless of the specification of continuation play). This shows that payoffs cannot sum to more than 3 with perfect monitoring.

In contrast, 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 can give him the necessary incentive to play M rather than L. In particular, suppose that player 1 mixes $\frac{1}{3}U + \frac{2}{3}D$ in period 1, and the monitoring structure is such that player 2 gets signal m ("play M") with probability 1 following (U, L), and gets signals m and r ("play R") with probability $\frac{1}{2}$ each following (D, L). Suppose further that player 1 plays T in period 2 if she played U in period 1, and plays B in period 2 if she played D in period 1. Then, when player 2 sees signal m in period 1, his posterior belief that player 1 played U in period 1 is

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

Player 2 therefore expects to face T and B in period 2 with probability $\frac{1}{2}$ each, so he is willing to play M rather than L. Meanwhile, player 1 is always rewarded with (T, M) in period 2 when she plays U in period 1, so she is willing to play U (as well as D) in period 1.

To summarize, the advantage of private monitoring is that pooling players' information sets (in this case, player 2's information sets after (U, L) and (D, L)) can make providing incentives easier.[7] A companion paper (Sugaya and Wolitzky 2017) develops this point in the context of some canonical models in industrial organization.

Below, we show that private monitoring cannot outperform mediated perfect monitoring when there exists a feasible payoff vector v such that no player i is tempted to deviate if she gets continuation payoff v_i when she conforms and is minmaxed when she deviates. This condition is violated in the current example because, when $\delta = \frac{1}{6}$, no feasible continuation payoff for player 2 is high enough to induce him to respond to T with M rather than L. Specifically, in the example the condition holds if and only if $\delta \geq \frac{19}{25}$.

4. Preliminary results about $E_{\text{med}}(\delta)$

We begin with two preliminary results about the equilibrium payoff set with mediated perfect monitoring. These results are important for both our result on when private monitoring cannot outperform mediated perfect monitoring (Theorem 1) and our characterization of the equilibrium payoff set with mediated perfect monitoring (Theorem 2).

Let $\underline{u}_i$ be player i's correlated minmax payoff, given by

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

Let $\alpha^*_{-i} \in \Delta(A_{-i})$ be a solution to this minmax problem. Let d_i be player i's greatest possible gain from a deviation at any recommendation profile, given by

$$d_i = \max_{r \in A, a_i \in A_i} u_i(a_i, r_{-i}) - u_i(r).$$

Let $\underline{w}_i$ be the lowest continuation payoff such that player i does not want to deviate at any recommendation profile when she is minmaxed forever if she deviates, given by

$$\underline{w}_i = \underline{u}_i + \frac{1 - \delta}{\delta} d_i.$$

Let

$$W_i = \{w \in \mathbb{R}^{|I|} : w_i \geq \underline{w}_i\}.$$

[7] As far as we know, the observation that players in a repeated game can benefit from imperfections in monitoring even in the presence of a mediator is original. Examples by Kandori (1991), Sekiguchi (2002), Mailath et al. (2002), and Miyahara and Sekiguchi (2013) show that players can benefit from imperfect monitoring in finitely repeated games (Kandori's example is described in Mailath and Samuelson (2006, Section 12.1.3)). However, in their examples this conclusion relies on the absence of a mediator, and is thus due to the possibilities for correlation opened up by private monitoring. The broader point that giving players more information can be bad for incentives is of course familiar.

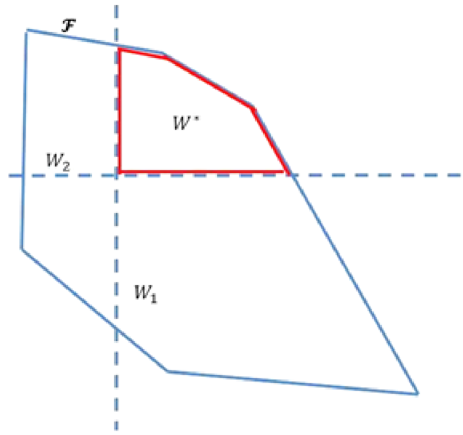

FIGURE 1. The intersection of $\mathcal{F}$, W_1, and W_2, denoted by W^*, satisfies $\mathring{W}^* \subseteq E_{\text{med}}(\delta)$.

Finally, let $\mathcal{F}$ be the convex hull of the set of feasible payoffs, let

$$W^* = \bigcap_{i \in I} W_i \cap \mathcal{F},$$

and denote the interior of W^* as a subspace of $\mathcal{F}$ by $\mathring{W}^*$.

Our first preliminary result is that all payoffs in $\mathring{W}^*$ are attainable in equilibrium with mediated perfect monitoring. See Figure 1. The intuition is that the mediator can virtually implement any payoff vector in W^* by minmaxing deviators.

We will actually prove the slightly stronger result that all payoffs in $\mathring{W}^*$ are attainable in a strict "full-support" equilibrium with mediated perfect monitoring. Formally, we say that an equilibrium has *full support* if for each player i and history $h_i^t = (r_{i,\tau}, a_\tau)_{\tau=1}^{t-1}$ such that there exist recommendations $(r_{-i,\tau})_{\tau=1}^{t-1}$ with $\Pr^\mu(r_\tau | (r_{\tau'}, a_{\tau'})_{\tau'=1}^{\tau-1}) > 0$ for each $\tau = 1, \ldots, t-1$, there exist alternative recommendations $(\bar{r}_{-i,\tau})_{\tau=1}^{t-1}$ such that for each $\tau = 1, \ldots, t-1$, we have

$$\Pr^\mu\big(r_{i,\tau}, \bar{r}_{-i,\tau} | (r_{i,\tau'}, \bar{r}_{-i,\tau'}, a_{\tau'})_{\tau'=1}^{\tau-1}\big) > 0 \quad \text{and} \quad \bar{r}_{-i,\tau} = a_{-i,\tau}.$$

That is, any history h_i^t consistent with the mediator's strategy is also consistent with i's opponents' equilibrium strategies (even if player i herself has deviated, noting that we allow $r_{i,\tau} \neq a_{i,\tau}$ in h_i^t). This is weaker than requiring that the mediator's recommendation has full support at all histories (on and off path), but stronger than requiring that the recommendation has full support at all on-path histories only. Note that if the equilibrium has full support, player i never believes that any of the other players has deviated.

LEMMA 1. *For all $v \in \mathring{W}^*$, there exists a strict full-support equilibrium with mediated perfect monitoring with payoff v. In particular, $\mathring{W}^* \subseteq E_{\text{med}}(\delta)$.*

PROOF. For each $v \in \mathring{W}^*$, there exists $\mu \in \Delta(A)$ such that $u(\mu) = v$ and $\mu(r) > 0$ for all $r \in A$. For each $i \in I$ and $\varepsilon \in (0, 1)$, approximate the minmax strategy α_{-i}^* by the full-support strategy $\alpha_{-i}^\varepsilon \equiv (1-\varepsilon)\alpha_{-i}^* + \varepsilon \sum_{a_{-i} \in A_{-i}} \frac{a_{-i}}{|A_{-i}|}$. Since $v \in \text{int}(\bigcap_{i \in I} W_i)$, there exists

$\varepsilon \in (0,1)$ such that, for each $i \in I$, we have

$$v_i > \max_{a_i \in A_i} u_i\big(a_i, \alpha_{-i}^{\varepsilon}\big) + \frac{1-\delta}{\delta} d_i. \tag{1}$$

Consider the following recommendation schedule: The mediator follows an automaton strategy whose state is identical to a subset of players $J \subseteq I$. Hence, the mediator has $2^{|I|}$ states. In the following construction of the mediator's strategy, J will represent the set of players who have ever deviated from the mediator's recommendation.

If the state J is equal to $\varnothing$ (no player has deviated), then the mediator recommends μ. If there exists i with $J = \{i\}$ (only player i has deviated), then the mediator recommends r_{-i} to players $-i$ according to $\alpha_{-i}^{\varepsilon}$, and recommends some best response to $\alpha_{-i}^{\varepsilon}$ to player i. Finally, if $|J| \geq 2$ (several players have deviated), then for each $i \in J$, the mediator recommends the best response to $\alpha_{-i}^{\varepsilon}$, while she recommends each profile $a_{-J} \in A_{-J}$ to the other players $-J$ with probability $\frac{1}{|A_{-J}|}$. The state transitions as follows: if the current state is J and players J' deviate, then the state transitions to $J \cup J'$.

Player i's strategy is to follow her recommendation $r_{i,t}$ in period t. She believes that the mediator's state is $\varnothing$ if she herself has never deviated, and believes that the state is $\{i\}$ if she has deviated.

Since the mediator's recommendation has full support, player i's belief is consistent. (In particular, no matter how many times player i has been instructed to minmax some player j, it is always infinitely more likely that these instructions resulted from randomization by the mediator rather than a deviation by player j.) If player i has deviated, then (given her belief) it is optimal for her to always play a static best response to $\alpha_{-i}^{\varepsilon}$, since the mediator always recommends $\alpha_{-i}^{\varepsilon}$ in state $\{i\}$. Given that a unilateral deviation by player i is punished in this way, (1) implies that on-path player i has a strict incentive to follow her recommendation $r_{i,t}$ at any recommendation profile $r_t \in A$. Hence, she has a strict incentive to follow her recommendation when she believes that $r_{-i,t}$ is distributed according to $\Pr^{\mu}(r_{-i,t}|h_i^t)$. □

The condition that $\mathring{W}^* \neq \varnothing$ can be more transparently stated as a lower bound on the discount factor. In particular, $\mathring{W}^* \neq \varnothing$ if and only if there exists $v \in \mathcal{F}$ such that

$$v_i > \underline{u}_i + \frac{1-\delta}{\delta} d_i \quad \text{for all } i \in I$$

or, equivalently,

$$\delta > \delta^* \equiv \min_{v \in \mathcal{F}} \max_{i \in I} \frac{d_i}{d_i + v_i - \underline{u}_i}. \tag{2}$$

For instance, it can be checked that $\delta^* = \frac{19}{25}$ in Example 1 of Section 3. Note that δ^* is strictly less than 1 if and only if the stage game admits a feasible and strictly individually rational payoff vector (relative to correlated minmax payoffs).[8] For most games of interest, δ^* will be some intermediate discount factor that is not especially close to either 0 or 1.

[8]Recall that a payoff vector v is *strictly individually rational* if $v_i > \underline{u}_i$ for all $i \in I$.

Our second preliminary result is that if a strict full-support equilibrium exists, then any payoff vector that can be attained by a mediator's strategy that is incentive compatible on path is (virtually) attainable in strict equilibrium.

Lemma 2. *With mediated perfect monitoring, fix a payoff vector v, and suppose there exists a mediator's strategy μ that (1) attains v when players obey the mediator, and (2) has the property that obeying the mediator is optimal for each player at each on-path history, when she is minmaxed forever if she deviates: that is, for each player i and on-path history h_m^{t+1},*

$$\begin{aligned}(1-\delta)\mathbb{E}\big[u_i(r_t) \mid h_m^t, r_{i,t}\big] + \delta\mathbb{E}\left[(1-\delta)\sum_{\tau=t+1}^{\infty}\delta^{\tau-t-1}u_i\big(\mu\big(h_m^\tau\big)\big)\,\middle|\, h_m^t, r_{i,t}\right] \\ \geq \max_{a_i\in A_i}(1-\delta)\mathbb{E}\big[u_i(a_i, r_{-i,t}) \mid h_m^t, r_{i,t}\big] + \delta\underline{u}_i.\end{aligned} \tag{3}$$

Suppose also that there exists a strict full-support equilibrium. (For example, such an equilibrium exists if $\mathring{W}^* \neq \emptyset$, by Lemma 1.) Then $v \in \overline{E_{\text{med}}(\delta)}$.[9]

Proof. Fix such a strategy μ and any strict full-support equilibrium μ^{strict}. We construct a strict equilibrium that attains a payoff close to v.

In period 1, the mediator draws one of two states, R_v and R_{perturb}, with probabilities $1-\varepsilon$ and ε, respectively. In state R_v, the mediator's recommendation is determined as follows: If no player has deviated up to period t, the mediator recommends r_t according to $\mu(h_m^t)$. If only player i has deviated, the mediator recommends $r_{-i,t}$ to players $-i$ according to α_{-i}^*, and recommends some best response to α_{-i}^* to player i. Multiple deviations are treated as in the proof of Lemma 1. Alternatively, in state R_{perturb}, the mediator follows the equilibrium μ^{strict}. Player i follows the recommendation $r_{i,t}$ in period t. Since the constructed recommendation schedule has full support, player i never believes that another player has deviated. Moreover, since μ^{strict} has full support, player i believes that the mediator's state is R_{perturb} with positive probability after any history. Therefore, by (3) and the fact that μ^{strict} is a strict equilibrium, it is always strictly optimal for each player i to follow her recommendation on path. Taking $\varepsilon \to 0$ yields a sequence of strict equilibria with payoffs converging to v. □

5. A sufficient condition for $\overline{E_{\text{med}}(\delta)}$ to give an upper bound

Our sufficient condition for mediated perfect monitoring to outperform private monitoring in two-player games is that $\delta > \delta^*$. In Section 9, we discuss what happens when there are more than two players or the condition that $\delta > \delta^*$ is relaxed.

Let $E(\delta, p)$ be the set of (possibly weak) sequential equilibrium payoffs with private monitoring structure p. Note that $E(\delta, p)$ is closed, as we use the product topology on assessments (Fudenberg and Levine 1983).

[9] Throughout, $\bar{X}$ denotes the closure of X.

Theorem 1. *If* $|I| = 2$ *and* $\delta > \delta^*$*, then for every private monitoring structure* p *and every nonnegative Pareto weight* $\lambda \in \Lambda_+ \equiv \{\lambda \in \mathbb{R}^2_+ : \|\lambda\| = 1\}$*, we have*

$$\max_{v \in E(\delta, p)} \lambda \cdot v \leq \max_{v \in \overline{E_{\text{med}}(\delta)}} \lambda \cdot v.$$

Theorem 1 says that in games involving two players of at least moderate patience, the Pareto frontier of the (closure of the strict) equilibrium payoff set with mediated perfect monitoring extends farther in any nonnegative direction than does the Pareto frontier of the equilibrium payoff set with any private monitoring structure.[10] We emphasize that Theorem 1 does not require that players are patient enough for the folk theorem to apply.

We describe the idea of the proof of Theorem 1, deferring the proof itself to Section 8.

Let $\mathcal{E}(\delta)$ be the equilibrium payoff set in the mediated repeated game with the following *universal monitoring structure*: the mediator directly observes the recommendation profile r_t and the action profile a_t in each period t, while each player i observes nothing beyond her own recommendation $r_{i,t}$ and her own action $a_{i,t}$.[11] This monitoring structure is so called because it embeds any private monitoring structure p by setting $\mu(h^t_m)$ equal to $p(\cdot|a_{t-1})$ for every history $h^t_m = (r_\tau, a_\tau)^{t-1}_{\tau=1}$.[12] It particular, we have $E(\delta, p) \subseteq \mathcal{E}(\delta)$ for every p, so to prove Theorem 1 it suffices to show that

$$\sup_{v \in \mathcal{E}(\delta)} \lambda \cdot v \leq \max_{v \in \overline{E_{\text{med}}(\delta)}} \lambda \cdot v. \tag{4}$$

To show this, the idea is to start with an equilibrium in $\mathcal{E}(\delta)$—where players only observe their own recommendations—and then show that the players' recommendations can be "publicized" without violating anyone's obedience constraints.[13] To see why this is possible (when $|I| = 2$ and $\delta > \delta^*$ or, equivalently, $\mathring{W}^* \neq \emptyset$), first note that we can restrict attention to equilibria with Pareto-efficient on-path continuation payoffs, as improving both players' on-path continuation payoffs improves their incentives (assuming that deviators are minmaxed, which is possible when $\mathring{W}^* \neq \emptyset$, by Lemma 2). Next, if $|I| = 2$ and $\mathring{W}^* \neq \emptyset$, then if a Pareto-efficient payoff vector v lies outside W_i for one player (say player 2), it must then lie inside W_j for the other player (player 1). Hence, at each history h^t, there can be only one player—here player 2—whose obedience constraint could be violated if we publicized both players' past recommendations.

Now suppose that at history h^t we do publicize the entire vector of players' past recommendations $r^t = (r_\tau)^{t-1}_{\tau=1}$, but the mediator then issues period t recommendations according to the original equilibrium distribution of recommendations conditional on *player* 2*'s past recommendations* $r^t_2 = (r_{2,\tau})^{t-1}_{\tau=1}$ *only*. We claim that doing this violates neither player's obedience constraint: Player 1's obedience constraint is easy to satisfy, as

[10]We do not know if the same result holds for negative Pareto weights.

[11]This information structure may not result from mediated communication among the players, as actions are not publicly observed. Again, we simply view $\mathcal{E}(\delta)$ as a technical device for bounding $E(\delta, p)$.

[12]Incidentally, this embedding does not yield an obedient equilibrium.

[13]More precisely, the construction in the proof both publicizes the players' recommendations and modifies the equilibrium in ways that only improve the players' λ-weighted payoffs.

we can always ensure that continuation payoffs lie in W_1, and since player 2 already knew r_2^t in the original equilibrium, publicizing h^t while issuing recommendations based only on r_2^t does not affect his incentives.

An important missing step in this proof sketch is that in the original equilibrium in $\mathcal{E}(\delta)$, at some histories it may be player 1 who is tempted to deviate when we publicize past recommendations, while it is player 2 who is tempted at other histories. For instance, it is not obvious how to publicize past recommendations when ex ante equilibrium payoffs are very good for player 1 (so player 2 is tempted to deviate in period 1), but continuation payoffs at some later history are very good for player 2 (so then player 1 is tempted to deviate). The proof of Theorem 1 shows that we can ignore this possibility, because—somewhat unexpectedly—equilibrium paths like this one are never needed to sustain Pareto-efficient payoffs. In particular, to sustain an ex ante payoff that is very good for player 1 (i.e., outside W_2), we never need to promise continuation payoffs that are very good for player 2 (i.e., outside W_1). The intuition is that, rather than promising player 2 a very good continuation payoff outside W_1, we can instead promise him a fairly good continuation inside W_1, while compensating him for this change by also occasionally transitioning to this fairly good continuation payoff at histories where the original promised continuation payoff is less good for him. Finally, since the feasible payoff set is convex, the resulting "compromise" continuation payoff vector is also acceptable to player 1.

A remark: The reader may wonder why we are not satisfied with simply bounding $E(\delta, p)$ by $\mathcal{E}(\delta)$. The answer is that the only way we know to recursively characterize the Pareto frontier of $\mathcal{E}(\delta)$ is by first establishing (4) and then characterizing the Pareto frontier of $\overline{E_{\text{med}}(\delta)}$. So this approach would not avoid the need to establish (4).

6. Recursively characterizing $\overline{E_{\text{med}}(\delta)}$

We have seen that $\overline{E_{\text{med}}(\delta)}$ is an upper bound on $E(\delta, p)$ for two-player games satisfying $\delta > \delta^*$. As our goal is to give a recursive upper bound on $E(\delta, p)$, it remains to recursively characterize $\overline{E_{\text{med}}(\delta)}$. Our characterization assumes that $\delta > \delta^*$, but it applies for any number of players.[14]

Recall that APS characterize the perfect public equilibrium set with imperfect public monitoring as the iterative limit of a generating operator B, where $B(W)$ is defined as the set of payoffs that can be sustained when on- and off-path continuation payoffs are drawn from W. We show that the sequential equilibrium payoff set with mediated perfect monitoring is the iterative limit of a generating operator $\tilde{B}$, where $\tilde{B}(W)$ is the set of payoffs that can be sustained when on-path continuation payoffs are drawn from W and deviators are minmaxed off path. There are two things to prove: (i) we can indeed minmax deviators off path, and (ii) on-path continuation payoffs must themselves be sequential equilibrium payoffs. The first of these facts is Lemma 2. For the second, note that in an obedient equilibrium with perfect monitoring, players can perfectly infer each

[14]The set $\overline{E_{\text{med}}(\delta)}$ admits a recursive characterization even if $\delta < \delta^*$, but in this case the characterization is somewhat more complicated. The details are available from the authors.

other's private history on path. Continuation play at on-path histories (but not off-path histories) is therefore common knowledge, which gives the desired recursive structure.

In what follows, we assume familiarity with APS and focus on the new features that emerge when mediation is available. Our terminology parallels that in Section 7.3 of Mailath and Samuelson (2006).

Definition 2. For any set $V \subseteq \mathbb{R}^{|I|}$, a correlated action profile $\alpha \in \Delta(A)$ is *minmax-threat enforceable* on V by a mapping $\gamma : A \to V$ if, for each player i and action $a_i \in \operatorname{supp} \alpha_i$,

$$\mathbb{E}^{\alpha}\big[(1-\delta)u_i(a_i, a_{-i}) + \delta\gamma(a_i, a_{-i})|a_i\big] \geq \max_{a_i' \in A_i} \mathbb{E}^{\alpha}\big[(1-\delta)u_i\big(a_i', a_{-i}\big)|a_i\big] + \delta\underline{u}_i.$$

Definition 3. A payoff vector $v \in \mathbb{R}^{|I|}$ is *minmax-threat decomposable* on V if there exists a correlated action profile $\alpha \in \Delta(A)$ that is minmax-threat enforced on V by a mapping γ such that

$$v = \mathbb{E}^{\alpha}\big[(1-\delta)u(a) + \delta\gamma(a)\big].$$

Let $\tilde{B}(V) = \{v \in \mathbb{R}^{|I|} : v \text{ is minmax-threat decomposable on } V\}$.

We show that the following algorithm recursively computes $\overline{E_{\text{med}}(\delta)}$: let $W^1 = \mathcal{F}$, $W^n = \tilde{B}(W^{n-1})$ for $n > 1$, and $W^\infty = \lim_{n\to\infty} W^n$.

Theorem 2. *If $\delta > \delta^*$, then $\overline{E_{\text{med}}(\delta)} = W^\infty$.*

With the exception of the following two lemmas, the proof of Theorem 2 is entirely standard and is omitted. The lemmas correspond to facts (i) and (ii) above. In particular, Lemma 3 follows directly from APS and Lemma 2, while Lemma 4 establishes on-path recursivity. For both lemmas, assume $\delta > \delta^*$.

Lemma 3. *If a set $V \subseteq \mathbb{R}^{|I|}$ is bounded and satisfies $V \subseteq \tilde{B}(V)$, then $\tilde{B}(V) \subseteq \overline{E_{\text{med}}(\delta)}$.*

Lemma 4. *We have $\overline{E_{\text{med}}(\delta)} = \tilde{B}(\overline{E_{\text{med}}(\delta)})$.*

Proof. By Lemma 3 and boundedness, we need only show that $\overline{E_{\text{med}}(\delta)} \subseteq \tilde{B}(\overline{E_{\text{med}}(\delta)})$.

Let $E_{\text{med}}^{\text{weak}}(\delta)$ be the set of (possibly weak) sequential equilibrium payoffs with mediated perfect monitoring. Note that in any sequential equilibrium, player i's continuation payoff at any history h_i^t must be at least $\underline{u}_i$. Therefore, if μ is an on-path recommendation strategy in a (possibly weak) sequential equilibrium, then it must satisfy (3). Hence, under the assumption that $\mathring{W}^* \neq \emptyset$, we have $E_{\text{med}}^{\text{weak}}(\delta) \subseteq \overline{E_{\text{med}}(\delta)}$.

Now, for any $v \in E_{\text{med}}(\delta)$, let μ be a corresponding equilibrium mediator's strategy. In the corresponding equilibrium, if some player i deviates in period 1 while her opponents are obedient, player i's continuation payoff must be at least $\underline{u}_i$. Hence, we have

$$\mathbb{E}^{\mu}\big[(1-\delta)u_i(a_i, a_{-i}) + \delta w_i(a_i, a_{-i})|a_i\big] \geq \max_{a_i' \in A_i} E^{\mu}\big[(1-\delta)u_i\big(a_i', a_{-i}\big)|a_i\big] + \delta\underline{u}_i,$$

where $w_i(a_i, a_{-i})$ is player i's equilibrium continuation payoff when action profile (a_i, a_{-i}) is recommended and obeyed in period 1. Finally, since action profile (a_i, a_{-i}) is in the support of the mediator's recommendation in period 1, each player assigns probability 1 to the true mediator's history when (a_i, a_{-i}) is recommended and played in period 1. Therefore, continuation play from this history is itself at least a weak sequential equilibrium. In particular, we have $w_i(a_i, a_{-i}) \in E_{\text{med}}^{\text{weak}}(\delta) \subseteq \overline{E_{\text{med}}(\delta)}$ for all $(a_i, a_{-i}) \in \operatorname{supp} \mu(h^t)$. Hence, v is minmax-threat decomposable on $\overline{E_{\text{med}}(\delta)}$ by action profile $\mu(\varnothing)$ and continuation payoff function w, so in particular $v \in \tilde{B}(\overline{E_{\text{med}}(\delta)})$.

We have shown that $E_{\text{med}}(\delta) \subseteq \tilde{B}(\overline{E_{\text{med}}(\delta)})$. As $\overline{E_{\text{med}}(\delta)}$ is compact and $\tilde{B}$ preserves compactness, taking closures yields $\overline{E_{\text{med}}(\delta)} \subseteq \overline{\tilde{B}(\overline{E_{\text{med}}(\delta)})} = \tilde{B}(\overline{E_{\text{med}}(\delta)})$. □

Combining Theorems 1 and 2 yields our main conclusion: *In two-player games with $\delta > \delta^*$, the equilibrium payoff set with mediated perfect monitoring is a recursive upper bound on the equilibrium payoff set with any imperfect private monitoring structure.*

7. The upper bound in an example

We illustrate our results with an application to a repeated Bertrand game. We compute the greatest equilibrium payoff that each firm can attain for any private monitoring structure.

Consider the following Bertrand game: There are two firms $i \in \{1, 2\}$, and each firm i's possible price level is $p_i \in \{W, L, M, H\}$ (price war, low price, medium price, high price). Given p_1 and p_2, firm i's profit is determined by the following payoff matrix:

	W	L	M	H
W	15, 15	30, 25	50, 15	80, 0
L	25, 30	40, 40	60, 35	90, 15
M	15, 50	35, 60	55, 55	85, 35
H	0, 80	15, 90	35, 85	65, 65

Example 2

Note that L (low price) is a dominant strategy in the stage game, W (price war) is a costly action that hurts the other firm, and (H, H) maximizes the sum of the firms' profits. The feasible payoff set is given by

$$\mathcal{F} = \text{co}\big\{(0, 80), (15, 15), (15, 90), (35, 85), (65, 65), (80, 0), (85, 35), (90, 15)\big\},$$

where we include only the extreme points in specifying the convex hull. In addition, each firm's minmax payoff $\underline{u}_i$ is 25, so the feasible and individually rational payoff set is given by

$$\text{co}\big\{(25, 25), (25, 87.5), (35, 85), (65, 65), (85, 35), (87.5, 25)\big\}.$$

In particular, the greatest feasible and individually rational payoff for each firm is 87.5. See Figure 2 for an illustration.

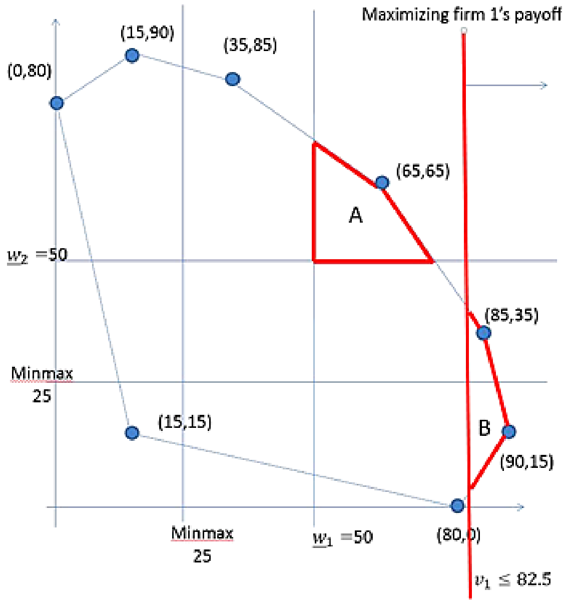

FIGURE 2. Region A is included in $E_{\text{med}}(\delta)$ by Lemma 1. Region B is not included in $E_{\text{med}}(\delta)$. Note that Region B includes payoffs which are feasible and individually rational.

In this game, each firm's maximum deviation gain d_i is 25. Since the game is symmetric, the critical discount factor δ^* above which we can apply Theorems 1 and 2 is given by plugging the best symmetric payoff of 65 into (2), which gives

$$\delta^* = \frac{25}{25 + 65 - 25} = \frac{5}{13}.$$

To illustrate our results, we find the greatest equilibrium payoff that each firm can attain for any private monitoring structure when $\delta = \frac{1}{2} > \frac{5}{13}$.

When $\delta = \frac{1}{2}$, we have

$$\underline{w}_i = 25 + \frac{1-\delta}{\delta} 25 = 50.$$

Hence, Lemma 1 implies that

$$\{v \in \mathring{\mathcal{F}} : v_i > 50 \text{ for each } i\} = \operatorname{int}\operatorname{co}\big\{(50,50),(75,50),(50,75),(65,65)\big\} \subseteq E_{\text{med}}(\delta).$$

We now compute the best payoff vector for firm 1 in $\tilde{B}(\mathcal{F})$. By Theorems 1 and 2, any Pareto-efficient payoff profile not included $\tilde{B}(\mathcal{F})$ is not included in $E(\delta, p)$ for any p.

In computing the best payoff vector for firm 1, it is natural to conjecture that firm 1's incentive compatibility constraint is not binding. We thus consider a relaxed problem with only firm 2's incentive constraint, and then verify that firm 1's incentive constraint is satisfied. Note that playing L is always the best deviation for firm 2. Furthermore, the corresponding deviation gain decreases as firm 1 increases its price from W to L, and (weakly) increases as it increases its price from L to M or H. In addition, firm 1's payoff increases as firm 1 increases its price from W to L and decreases as it increases its price from L to M or H. Hence, so as to maximize firm 1's payoff, firm 1 should play L.

Suppose that firm 2 plays H. Then firm 2's incentive compatibility constraint is

$$(1-\delta)\underbrace{25}_{\text{maximum deviation gain}} \le \delta(w_2 - \underbrace{25}_{\text{minmax payoff}}),$$

where w_2 is firm 2's continuation payoff; that is, $w_2 \ge 50$.

By feasibility, $w_2 \ge 50$ implies that $w_1 \le 75$. Hence, if $r_2 = H$, the best minmax-threat decomposable payoff for firm 1 is

$$(1-\delta)\begin{pmatrix}90\\15\end{pmatrix} + \delta\begin{pmatrix}75\\50\end{pmatrix} = \begin{pmatrix}82.5\\32.5\end{pmatrix}.$$

Since 82.5 is larger than any payoff that firm 1 can get when firm 2 plays W, M, or L, firm 2 should indeed play H to maximize firm 1's payoff. Moreover, since $75 \ge \underline{w}_1$, firm 1's incentive constraint is not binding. Thus, we have shown that 82.5 is the best payoff for firm 1 in $\tilde{B}(\mathcal{F})$.

In contrast, with mediated perfect monitoring it is in fact possible to (virtually) implement an action path in which firm 1's payoff is 82.5: play (L, H) in period 1 (with payoffs $(90, 15)$) and then play $\frac{1}{2}(M, H) + \frac{1}{2}(H, H)$ forever (with payoffs $\frac{1}{2}(85, 35) + \frac{1}{2}(65, 65) = (75, 50)$), while minmaxing deviators.

Thus, when $\delta = \frac{1}{2}$, each firm's greatest feasible and individually rational payoff is 87.5, but the greatest payoff it can attain with any imperfect private monitoring structure is only 82.5. In this simple game, we can therefore say exactly how much of a constraint is imposed on each firm's greatest equilibrium payoff by the firms' impatience alone, independently of the monitoring structure.

8. Proof of Theorem 1

8.1 *Preliminaries and plan of proof*

We wish to establish (4) for every Pareto weight $\lambda \in \Lambda_+$. As $\mathcal{E}(\delta)$ is convex, it suffices to establish

$$\max_{v \in \mathcal{E}_0(\delta)} \lambda \cdot v \le \max_{v \in \overline{E_{\text{med}}(\delta)}} \lambda \cdot v$$

for every compact set $\mathcal{E}_0(\delta) \subseteq \mathcal{E}(\delta)$.

Fix a compact set $\mathcal{E}_0(\delta) \subseteq \mathcal{E}(\delta)$. Note that Lemma 1 implies that

$$W^* \subseteq \overline{E_{\text{med}}(\delta)}.$$

Therefore, for every Pareto weight $\lambda \in \Lambda_+$, if there exists $v \in \arg\max_{v' \in \mathcal{E}_0(\delta)} \lambda \cdot v'$ such that $v \in W^*$, then there exists $v^* \in \overline{E_{\text{med}}(\delta)}$ such that $\lambda \cdot v \le \lambda \cdot v^*$, as desired.

Hence, we are left to consider $\lambda \in \Lambda_+$ with

$$\arg\max_{v' \in \mathcal{E}_0(\delta)} \lambda \cdot v' \cap W^* = \emptyset. \tag{5}$$

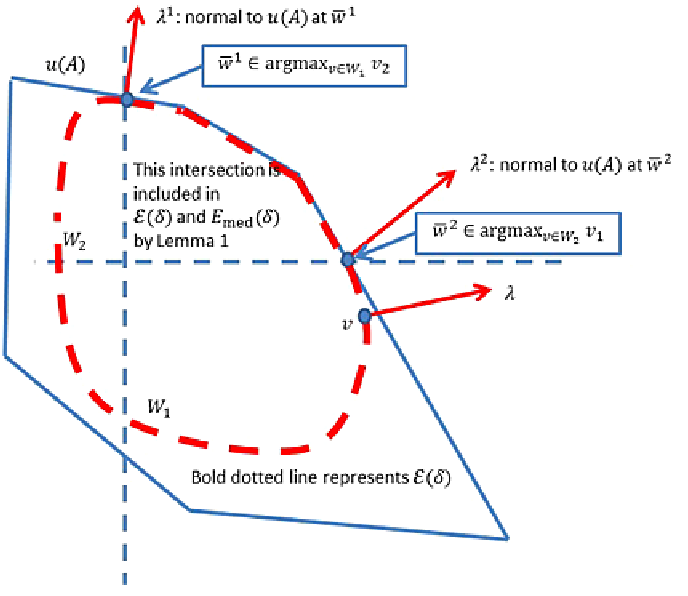

FIGURE 3. Setup for the construction. The green region (intersection of $u(A)$, W_1, and W_2) is included in $E_{\text{med}}(\delta)$ and $\mathcal{E}(\delta)$ by Lemma 1. Bold dotted line represents $\mathcal{E}(\delta)$.

Since we consider two-player games, we can order $\lambda \in \Lambda_+$ as $\lambda \leq \lambda'$ if and only if $\frac{\lambda_1}{\lambda_2} \leq \frac{\lambda'_1}{\lambda'_2}$, that is, the vector λ is steeper than λ'. For each player i, let $\bar{w}^i$ be the Pareto-efficient point in W_i satisfying

$$\bar{w}^i \in \arg\max_{v \in W_i \cap \mathcal{F}} v_{-i}.$$

Note that the assumption that $\mathring{W}^* \neq \varnothing$ implies that $\bar{w}^i \in W^*$. Let $\alpha^i \in \Delta(A)$ be a recommendation that attains $\bar{w}^i$: $u(\alpha^i) = \bar{w}^i$. Let Λ^i be the (nonempty) set of Pareto weight λ^i such that $\bar{w}^i \in \arg\max_{v \in \mathcal{F}} \lambda^i \cdot v$:

$$\Lambda^i = \left\{\lambda^i \in \mathbb{R}^2_+ : \|\lambda^i\| = 1, \bar{w}^i \in \arg\max_{v \in \mathcal{F}} \lambda^i \cdot v\right\}.$$

As $\mathcal{F}$ is convex, if λ satisfies (5), then either $\lambda < \lambda^1$ for each $\lambda^1 \in \Lambda^1$ or $\lambda > \lambda^2$ for each $\lambda^2 \in \Lambda^2$. See Figure 3. We focus on the case where $\lambda > \lambda^2$. (The proof for the $\lambda < \lambda^1$ case is symmetric and thus is omitted.)

Fix $v \in \arg\max_{v' \in \mathcal{E}_0(\delta)} \lambda \cdot v'$. Let $(\mu, (\sigma_i)_{i \in I})$ be an equilibrium that attains v with the universal monitoring structure (where players do not observe each other's actions). By Lemma 2, it suffices to construct a mediator's strategy μ^* yielding payoffs v^* such that (3) (perfect monitoring incentive compatibility) holds and $\lambda \cdot v \leq \lambda \cdot v^*$. The rest of the proof constructs such a strategy.

The plan for constructing the strategy μ^* is as follows: First, from μ we construct a mediator's strategy $\bar{\mu}$ that yields payoffs v and satisfies perfect monitoring incentive compatibility for player 2, but possibly not for player 1. The idea is to set the distribution of recommendations under $\bar{\mu}$ equal to the distribution of recommendations under

μ conditional on player 2's information only. Second, from $\bar{\mu}$, we construct a mediator's strategy μ^* that yields payoffs v^* with $\lambda \cdot v \leq \lambda \cdot v^*$ and satisfies perfect monitoring incentive compatibility for both players.

8.2 *Construction and properties of* $\bar{\mu}$

For each on-path history of player 2's recommendations, denoted by $r_2^t = (r_{2,\tau})_{\tau=1}^{t-1}$ (with $r_2^1 = \{\varnothing\}$), let $\Pr^\mu(\cdot|r_2^t)$ be the conditional distribution of recommendations in period t, and let $w^\mu(r_2^t)$ be the continuation payoff vector from period t onward conditional on r_2^t:

$$w^\mu\left(r_2^t\right) = \mathbb{E}^\mu\left[\sum_{\tau=0}^{\infty} \delta^\tau u(r_{t+\tau}) \,\middle|\, r_2^t\right].$$

Define $\bar{\mu}$ so that, for every on-path history $r^t = (r_\tau)_{\tau=1}^{t-1}$ (with $r^1 = \{\varnothing\}$), the mediator draws r_t according to $\Pr^\mu(r_t|r_2^t)$:

$$\Pr^{\bar{\mu}}\left(r_t|r^t\right) \equiv \Pr^\mu\left(r_t|r_2^t\right).$$

We claim that $\bar{\mu}$ yields payoffs v and satisfies (3) for player 2. To see this, let $w^{\bar{\mu}}(r^t)$ be the continuation payoff vector from period t onward conditional on r^t under $\bar{\mu}$, and note that $w^{\bar{\mu}}(r^t) = w^\mu(r_2^t)$. In particular, $w^{\bar{\mu}}(r^1) = w^\mu(r_2^1) = v$ since $r^1 = r_2^1 = \{\varnothing\}$. In addition, the fact that μ is an equilibrium with the universal monitoring structure implies that, for every on-path history r^{t+1},

$$(1-\delta)\mathbb{E}^\mu\left[u_2(r_t)|r_2^{t+1}\right] + \delta w^\mu\left(r_2^{t+1}\right) \geq \max_{a_2 \in A_2} (1-\delta)\mathbb{E}^\mu\left[u_2(r_{1,t}, a_2)|r_2^{t+1}\right] + \delta \underline{u}_2.$$

As $w_2^\mu(r_2^{t+1}) = w_2^{\bar{\mu}}(r^{t+1})$ and $\Pr^\mu(r_t|r_2^{t+1}) = \Pr^{\bar{\mu}}(r_t|r^t, r_{2,t})$, this implies that (3) holds for player 2.

8.3 *Construction of* μ^*

The mediator's strategy μ^* will involve mixing over continuation payoffs at certain histories r^{t+1}, and we will denote the mixing probability at history r^{t+1} by $\rho(r^{t+1})$. Our approach is to first construct the mediator's strategy μ^* for an arbitrary function $\rho : \bigcup_{t=1}^\infty A^{t-1} \to [0,1]$ specifying these mixing probabilities, and to then specify the function ρ.

Given a function $\rho : \bigcup_{t=1}^\infty A^{t-1} \to [0,1]$, the mediator's strategy μ^* is defined as follows: In each period $t = 0, 1, 2, \ldots$, the mediator is in one of two states, $\omega_t \in \{S_1, S_2\}$ (where period 0 is a purely notational, and as usual the game begins in period 1). Given the state, recommendations in period $t \geq 1$ are as follows:

(i) In state S_1, at history $r^t = (r_\tau)_{\tau=1}^{t-1}$, the mediator recommends r_t according to $\Pr^{\bar{\mu}}(r_t|r^t)$.

(ii) In state S_2, the mediator recommends r_t according to some $\alpha^1 \in \Delta(A)$ such that $u(\alpha^1) = \bar{w}^1$.

The initial state is $\omega_0 = S_1$. State S_2 is absorbing: if $\omega_t = S_2$, then $\omega_{t+1} = S_2$. Finally, the transition rule in state S_1 is as follows:

(i) If $w^{\bar{\mu}}(r^{t+1}) \notin W_1$, then $\omega_{t+1} = S_2$ with probability 1.

(ii) If $w^{\bar{\mu}}(r^{t+1}) \in W_1$, then $\omega_{t+1} = S_2$ with probability $1 - \rho(r^{t+1})$.

Thus, strategy μ^* agrees with $\bar{\mu}$, with the exception that μ^* occasionally transitions to an absorbing state where actions yielding payoffs $\bar{w}^1$ are recommended forever. In particular, such a transition always occurs when continuation payoffs under $\bar{\mu}$ lie outside W_1, and otherwise this transition occurs with probability $1 - \rho(r^{t+1})$.

To complete the construction of μ^*, it remains only to specify the function ρ. To this end, it is useful to define an operator F, which maps functions $w : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}^2$ to functions $F(w_2) : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}^2$. The operator F will be defined so that its unique fixed point is precisely the continuation value function in state S_1 under μ^* for a particular function ρ, and this function will be the one we use to complete the construction of μ^*.

Given $w : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}^2$, define $w^*(w) : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}$ so that, for every $r^t \in A^{t-1}$, we have

$$w^*(w)(r^t) = (1-\delta)u(\bar{\mu}(r^t)) + \delta\mathbb{E}[w(r^{t+1})|r^t]. \tag{6}$$

Next, given $w^*(w) : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}$, define $F(w) : \bigcup_{t=1}^{\infty} A^{t-1} \to \mathbb{R}$ so that, for every $r^t \in A^{t-1}$, we have

$$F(w)(r^t) = 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}} \left\{ \begin{array}{l} \rho(w)(r^t) \times w^*(w)(r^t) \\ \quad + (1-\rho(w)(r^t)) \times \bar{w}^1 \end{array} \right\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}} \bar{w}^1, \tag{7}$$

where, when $w^{\bar{\mu}}(r^t) \in W_1$, $\rho(w)(r^t)$ is the largest number in $[0, 1]$ such that

$$\rho(w)(r^t) \times w_2^*(w)(r^t) + (1-\rho(w)(r^t)) \times \bar{w}_2^1 \geq w_2^{\bar{\mu}}(r^t). \tag{8}$$

That is, if $w_2^*(w)(r^t) \geq w_2^{\bar{\mu}}(r^t)$, then $\rho(w)(r^t) = 1$, and otherwise, since $w^{\bar{\mu}}(r^t) \in W_1$ implies that $w_2^{\bar{\mu}}(r^t) \leq \bar{w}_2^1$, $\rho(w)(r^t) \in [0, 1]$ solves

$$\rho(w)(r^t) \times w_2^*(w)(r^t) + (1-\rho(w)(r^t)) \times \bar{w}_2^1 = w_2^{\bar{\mu}}(r^t).$$

(Intuitively, the term $1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1$ in (7) reflects the fact that we have replaced continuation payoffs outside of W_1 with player 2's most favorable continuation payoff within W_1, namely $\bar{w}^1$. This replacement may reduce player 2's value below his original value of $w_2^{\bar{\mu}}(r^t)$. However, (8) ensures that by also replacing continuation payoffs *within* W_1 with $\bar{w}^1$ with high enough probability, player 2's value does not fall below $w_2^{\bar{\mu}}(r^t)$.)

To show that F has a unique fixed point, it suffices to show that F is a contraction.

Lemma 5. *For all w and $\tilde{w}$, we have $\|F(w) - F(\tilde{w})\| \leq \delta\|w - \tilde{w}\|$, where $\|w - \tilde{w}\| \equiv \sup_{r^t} \|w(r^t) - \tilde{w}(r^t)\|$.*

Proof. By (6), $\|w^*(w) - w^*(\tilde{w})\| \leq \delta\|w - \tilde{w}\|$. By (7),

$$\big|F(w)\big(r^t\big) - F(\tilde{w})\big(r^t\big)\big| = 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}} \left| \begin{array}{l} \big\{\rho(w)\big(r^t\big)w^*(w)\big(r^t\big) + \big(1-\rho(w)\big(r^t\big)\big)\bar{w}^1\big\} \\ \quad - \big\{\rho(\tilde{w})\big(r^t\big)w^*(\tilde{w})\big(r^t\big) + \big(1-\rho(\tilde{w})\big(r^t\big)\big)\bar{w}^1\big\} \end{array} \right|$$
$$\leq \big\|w^*(w) - w^*(\tilde{w})\big\|.$$

Combining these inequalities yields $\|F(w) - F(\tilde{w})\| \leq \delta\|w - \tilde{w}\|$. □

Let w be the unique fixed point of F. Given this function w, let $w^* = w^*(w)$ (given by (6)) and let $\rho = \rho(w)$ (given by (8)). This completes the construction of the mediator's strategy μ^*.

8.4 *Properties of μ^**

Observe that

$$w^*\big(r^t\big) = (1-\delta)u\big(\bar{\mu}\big(r^t\big)\big) + \delta\mathbb{E}\big[w\big(r^{t+1}\big)|r^t\big] \tag{9}$$

and

$$w\big(r^t\big) = 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}}\big\{\rho\big(r^t\big)w^*\big(r^t\big) + \big(1-\rho\big(r^t\big)\big)\bar{w}^1\big\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1. \tag{10}$$

Thus, for $i = 1, 2$, $w_i^*(r^t)$ is player i's expected continuation payoff from period t given r^t and $\omega_t = S_1$ (before she observes $r_{i,t}$), and $w_i(r^t)$ is player i's expected continuation payoff from period t given r^t and $\omega_{t-1} = S_1$ (before she observes $r_{i,t}$). In particular, recalling that $\omega_1 = S_1$ and $v = w^{\mu}(\varnothing) \in W_1$, (8) implies that the ex ante payoff vector v^* is given by

$$v^* = w(\varnothing) = \rho(\varnothing)w^*(\varnothing) + \big(1-\rho(\varnothing)\big)\bar{w}^1.$$

We prove the following key lemma in Appendix B.

Lemma 6. *For all $t \geq 1$, if $w^{\bar{\mu}}(r^t) \in W_1$, then $\rho(r^t)w^*(r^t) + (1-\rho(r_2^t))\bar{w}^1$ Pareto dominates $w^{\bar{\mu}}(r^t)$.*

Here is a graphical explanation of Lemma 6: By (9), $w^*(r^t) - w^{\bar{\mu}}(r^t)$ is parallel to $w(r^{t+1}) - w^{\bar{\mu}}(r^{t+1})$. To evaluate this difference, consider (10) for period $t+1$. The term $1_{\{w^{\bar{\mu}}(r^{t+1})\notin W_1\}}\bar{w}^1$ indicates that we construct $w(r^{t+1})$ by replacing some continuation payoff not included in W_1 with $\bar{w}^1$. Hence, $w(r^{t+1}) - w^{\bar{\mu}}(r^{t+1})$ (and thus $w^*(r^t) - w^{\bar{\mu}}(r^t)$) is parallel to $\bar{w}^1 - \hat{w}(r^{t+1})$ for some $\hat{w}(r^{t+1}) \in \mathcal{F} \setminus W_1$. See Figure 4 for an illustration.

Recall that $\rho(r^t)$ is determined by (8). Since the vector $w^*(r^t) - w^{\bar{\mu}}(r^t)$ is parallel to $\bar{w}^1 - \hat{w}(r^{t+1})$ for some $\hat{w}(r^{t+1}) \in \mathcal{F} \setminus W_1$ and $\mathcal{F}$ is convex, we have $w_1^*(r^t) \geq w_1^{\bar{\mu}}(r^t)$. Hence, if we take $\rho(r^t)$ so that the convex combination of $w_2^*(r^t)$ and $\bar{w}_2^1$ is equal to $w_2^{\bar{\mu}}(r^t)$, then player 1 is better off compared to $w_1^{\bar{\mu}}(r^t)$. See Figure 5.

Given Lemma 6, we show that μ^* satisfies perfect monitoring incentive compatibility (3) for both players, and $\lambda \cdot v \leq \lambda \cdot v^*$.

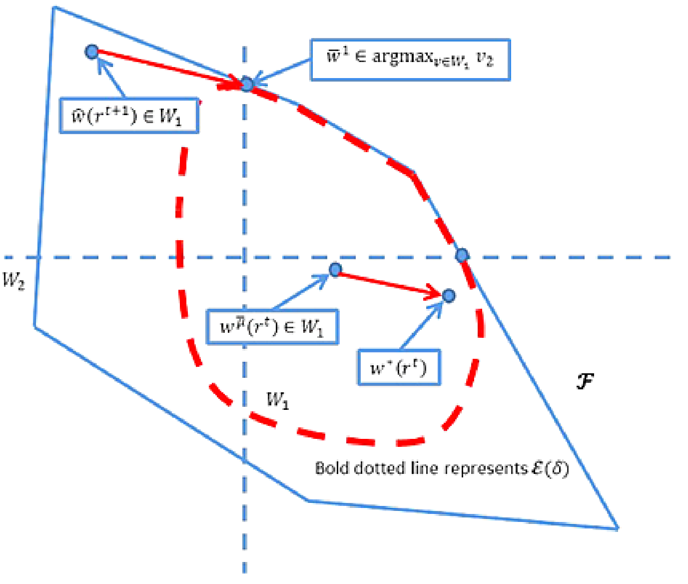

FIGURE 4. The vector from $w^{\bar{\mu}}(r^t)$ to $w^*(r^t)$ is parallel to the one from $\hat{w}(r^{t+1})$ to $\bar{w}^1$.

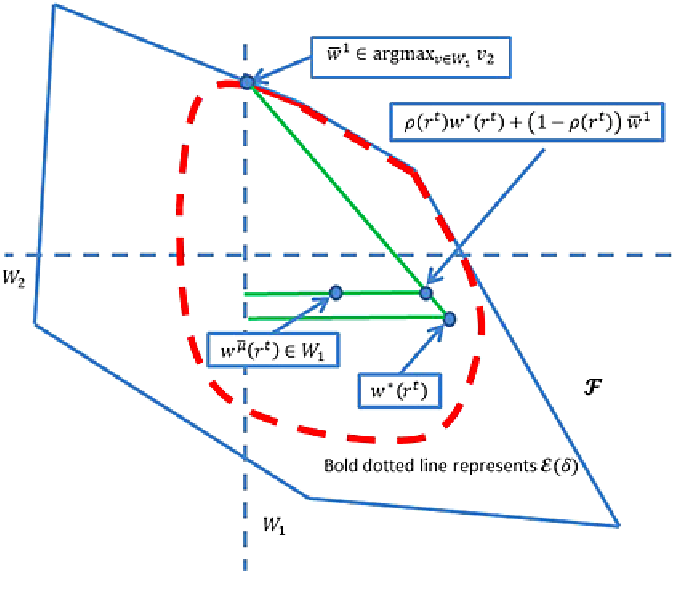

FIGURE 5. $\rho(r^t)w^*(r^t) + (1 - \rho(r^t))\bar{w}^1$ and $w^{\bar{\mu}}(r^t)$ have the same value for player 2.

(i) Incentive compatibility for player 1. It suffices to show that, conditional on any on-path history r^t and period t recommendation $r_{1,t}$, the expected continuation payoff from period $t+1$ onward lies in W_1. If $\omega_t = S_2$, then this continuation payoff is $\bar{w}^1 \in W_1$. If $\omega_t = S_1$, then it suffices to show that $w(r^{t+1}) \in W_1$ for all r^{t+1}. If $w^{\bar{\mu}}(r^{t+1}) \in W_1$, then, by Lemma 6, $w(r^{t+1}) = \rho(r^{t+1})w^*(r^{t+1}) + (1 - \rho(r^{t+1}))\bar{w}^1$ Pareto dominates $w^{\bar{\mu}}(r^{t+1}) \in W_1$, so $w(r^{t+1}) \in W_1$. If $w^{\bar{\mu}}(r^{t+1}) \notin W_1$, then $w(r^{t+1}) = \bar{w}^1 \in W_1$. Hence, $w(r^{t+1}) \in W_1$ for all r^{t+1}.

(ii) Incentive compatibility for player 2. Fix an on-path history r^t and a period t recommendation $r_{2,t}$. If $\omega_t = S_2$ or if both $\omega_t = S_1$ and $w^{\bar{\mu}}(r^{t+1}) \notin W_1$, then the expected continuation payoff from period $t+1$ onward conditional on $(r^t, r_{2,t})$ is $\bar{w}^1 \in W_1$, so (3) holds. If instead $\omega_t = S_1$ and $w^{\bar{\mu}}(r^{t+1}) \in W_1$, then $w(r^{t+1}) = \rho(r^{t+1})w^*(r^{t+1}) + (1-\rho(r^{t+1}))\bar{w}^1 = w^{\bar{\mu}}(r^{t+1})$ by (8). As μ is an equilibrium with the universal monitoring structure and $\Pr^{\mu^*}(r_t|r_2^{t+1}) = \Pr^{\bar{\mu}}(r_t|r^t, r_{2,t})$, this implies that (3) holds for player 2, by the same argument as in Section 8.2.

(iii) That $\lambda \cdot v \leq \lambda \cdot v^*$ is immediate from Lemma 6 with $t = 1$.

9. Extensions

This section discusses the extent to which the payoff bound is tight, as well as what happens when the conditions for Theorem 1 are violated.

9.1 *Tightness of the bound*

There are at least two senses in which $\overline{E_{\text{med}}(\delta)}$ is a *tight* bound on the equilibrium payoff set from the perspective of an observer who does not know the monitoring structure.

First, thus far our model of repeated games with private monitoring has maintained the standard assumption that the distribution of period t signals depends only on period t actions: that is, that this distribution can be written as $p(\cdot|a_t)$. In many settings, it would be desirable to relax this assumption and let the distribution of period t signals depend on the entire history of actions and signals up to period t, leading to a conditional distribution of the form $p_t(\cdot|a^t, z^t)$, as well as letting players receive signals before the first round of play. (Recall that $a^t = (a_\tau)_{\tau=1}^{t-1}$ and $z^t = (z_\tau)_{\tau=1}^{t-1}$.) For example, colluding firms do not only observe their sales in every period, but also occasionally get more information about their competitors' past behavior from trade associations, auditors, tax data, and the like.[15] From the perspective of an observer who finds such *nonstationary* private monitoring structures possible, the bound $\overline{E_{\text{med}}(\delta)}$ is clearly tight: $\overline{E_{\text{med}}(\delta)}$ is an upper bound on $E(\delta, p)$ for any nonstationary private monitoring structure p, because the equilibrium payoff set with the universal monitoring structure, $\mathcal{E}(\delta)$, remains an upper bound on $E(\delta, p)$, and the bound is tight because perfect monitoring with a given strategy of the mediator's itself induces a particular nonstationary private monitoring structure.

Second, from the perspective of an observer who finds only stationary monitoring structures possible, the bound $\overline{E_{\text{med}}(\delta)}$ is tight if the players can communicate through cheap talk, as they can then "replicate" the mediator among themselves. For this result, we also need to slightly generalize our definition of a private monitoring structure by letting the players receive signals before the first round of play, so that these signals can be used as a correlating device. This seems innocuous, especially if we take the perspective of an an outside observer who does not know the game's start date. The monitoring

[15]Rahman (2014, p. 1) quotes from the European Commission decision on the amino acid cartel: a typical cartel member "reported its citric acid sales every month to a trade association, and every year, Swiss accountants audited those figures."

structure is required to be stationary thereafter.[16] We call such a monitoring structure a *private monitoring structure with ex ante correlation*. Let $E_{\text{talk}}(\delta, p)$ be the sequential equilibrium payoff set in the repeated game with private monitoring structure with ex ante correlation p and finitely many rounds of public cheap talk before each round of play.

Proposition 2. *If $|I| = 2$ and $\delta > \delta^*$, then there exists a private monitoring structure with ex ante correlation p such that $\overline{E_{\text{talk}}(\delta, p)} = \overline{E_{\text{med}}(\delta)}$.*

The proof is long and is deferred to the Supplement. The main idea is as in the literature on implementing correlated equilibria without a mediator (see Forges (2009) for a survey). More specifically, Proposition 2 is similar to Theorem 9 of Heller et al. (2012), which shows that communication equilibria in repeated games with perfect monitoring can always be implemented by ex ante correlation and cheap talk. Since we also assume players observe actions perfectly, the main difference between the results is that theirs is for Nash rather than sequential equilibrium, so they are concerned only with detecting deviations rather than providing incentives to punish deviations once detected. In our model, when $\delta > \delta^*$, incentives to minmax the deviator can be provided (as in Lemma 1) *if* her opponent does not realize that the punishment phase has begun. The additional challenge in the proof of Proposition 2 is thus that we sometimes need a player to switch to the punishment phase for her opponent without realizing that this switch has occurred.

If one insists on stationary monitoring and does not allow communication, we believe that there are some games in which our bound is not tight, in that there are points in $\overline{E_{\text{med}}(\delta)}$ that are not attainable in equilibrium for any stationary private monitoring structure. We leave this as a conjecture.[17]

9.2 *What if $\mathring{W}^* = \emptyset$?*

The assumption that $\mathring{W}^* \neq \emptyset$ guarantees that all action profiles are supportable in equilibrium, which plays a key role in our results. However, this assumption is restrictive, in that it is violated when players are too impatient. Furthermore, it implies that the Pareto frontier of $\overline{E_{\text{med}}(\delta)}$ coincides with the Pareto frontier of the feasible payoff set for some Pareto weights λ (but of course not for others), so this assumption must also be relaxed for our approach to be able to give nontrivial payoff bounds for all Pareto weights.

To address these concerns, this subsection shows that even if $\mathring{W}^* = \emptyset$, $\overline{E_{\text{med}}(\delta)}$ may still be an upper bound on $E(\delta, p)$ for any private monitoring structure p, and $\overline{E_{\text{med}}(\delta)}$

[16] In particular, the distribution of signals can be the same every period. All we require is that the first of these signals arrives before the first round of play.

[17] Strictly speaking, since our maintained definition of a private monitoring structure does not allow ex ante correlation, if $\delta = 0$, then there are points in $\overline{E_{\text{med}}(\delta)}$ that are not attainable with any private monitoring structure whenever the stage game's correlated equilibrium payoff set strictly contains its Nash equilibrium payoff set. The nontrivial conjecture is that the bound is still not tight when ex ante correlation is allowed, but communication is not.

can still be characterized recursively. The idea is that even if not all action profiles are supportable, our approach still applies if a condition analogous to $\mathring{W}^* \neq \varnothing$ holds with respect to the subset of action profiles that are supportable.

Let $\operatorname{supp}(\delta)$ be the set of supportable actions with the universal monitoring structure:

$$\operatorname{supp}(\delta) = \left\{ a \in A : \begin{array}{l} \text{with the universal monitoring structure,} \\ \text{there exist an equilibrium strategy } \mu \\ \text{and history } h_m^t \text{ with } a \in \operatorname{supp}\bigl(\mu\bigl(h_m^t\bigr)\bigr) \end{array} \right\}.$$

Note that in this definition h_m^t can be an off-path history.

Next, given a product set of action profiles $\bar{A} = \prod_{i \in I} \bar{A}_i \subseteq A$, let $S_i(\bar{A})$ be the set of actions $a_i \in \bar{A}_i$ such that there exists a correlated action $\alpha_{-i} \in \Delta(\bar{A}_{-i})$ with

$$\begin{aligned} &(1-\delta) u_i(a_i, \alpha_{-i}) + \delta \max_{\bar{a} \in \bar{A}} u_i(\bar{a}) \\ &\quad \geq (1-\delta) \max_{\hat{a}_i \in A_i} u_i(\hat{a}_i, \alpha_{-i}) + \delta \min_{\hat{\alpha}_{-i} \in \Delta(\bar{A}_{-i})} \max_{a_i \in A_i} u_i(a_i, \hat{\alpha}_{-i}). \end{aligned}$$

That is, $a_i \in S_i(\bar{A})$ if there exists $\alpha_{-i} \in \Delta(\bar{A}_{-i})$ such that if her opponents play α_{-i}, player i's reward for playing a_i is the best payoff possible among those with support in $\bar{A}$, and player i's punishment for deviating from a_i is the worst possible among those with support in $\bar{A}_{-i}$, then player i plays a_i. Let $S(\bar{A}) = \prod_{i \in I} S_i(A_i) \subseteq \bar{A}$. Let $\mathcal{A}^1 = A$, let $\mathcal{A}^n = S(\mathcal{A}^{n-1})$ for $n > 1$, and let $\mathcal{A}^\infty = \lim_{n \to \infty} \mathcal{A}^n$. Note that the problem of computing $\mathcal{A}^\infty$ is tractable, as the set $S(\bar{A})$ is defined by a finite number of linear inequalities.

Finally, in analogy with the definition of w_i from Section 4, note that

$$\min_{\alpha_{-i} \in \Delta(\mathcal{A}_{-i}^\infty)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) + \frac{1-\delta}{\delta} \max_{r \in \mathcal{A}^\infty, a_i \in A_i} \bigl\{ u_i(a_i, r_{-i}) - u_i(r) \bigr\}$$

is the lowest continuation payoff such that player i does not want to deviate to any $a_i \in A_i$ at any recommendation profile $r \in \mathcal{A}^\infty$, when she is minmaxed forever if she deviates, *subject to the constraint that punishments are drawn from* $\mathcal{A}^\infty$. In analogy with the definition of W_i from Section 4, let

$$\begin{aligned} \bar{W}_i = \biggl\{ w \in \mathbb{R}^{|I|} : w_i \geq{} & \min_{\alpha_{-i} \in \Delta(\mathcal{A}_{-i}^\infty)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) \\ & + \frac{1-\delta}{\delta} \max_{r \in \mathcal{A}^\infty, a_i \in A_i} \bigl\{ u_i(a_i, r_{-i}) - u_i(r) \bigr\} \biggr\}. \end{aligned}$$

Proposition 3. *Assume that* $|I| = 2$. *If*

$$\operatorname{int}\Bigl(\bigcap_{i \in I} \bar{W}_i \cap \operatorname{co} u\bigl(\mathcal{A}^\infty\bigr) \Bigr) \neq \varnothing \tag{11}$$

in the topology induced from $\operatorname{co} u(\mathcal{A}^\infty)$*, then for every private monitoring structure* p *and every nonnegative Pareto weight* $\lambda \in \Lambda_+$*, we have*

$$\max_{v\in E(\delta,p)} \lambda\cdot v \leq \max_{v\in \overline{E_{\text{med}}(\delta)}} \lambda\cdot v.$$

In addition, $\operatorname{supp}(\delta) = \mathcal{A}^\infty$.

Proof. We show that $\operatorname{supp}(\delta) = \mathcal{A}^\infty$ whenever $\operatorname{int}(\bigcap_{i\in I} \bar{W}_i \cap \operatorname{co} u(\mathcal{A}^\infty)) \neq \varnothing$. Given this, the proof is analogous to the proof of Theorem 1, everywhere replacing $\mathcal{F} = \operatorname{co} u(A)$ with $\operatorname{co} u(\mathcal{A}^\infty)$ and replacing "full support" with "full support within $\mathcal{A}^\infty$."

We first show that $\operatorname{supp}(\delta) \subseteq \mathcal{A}^\infty$. For each i and n, we show that any action $a_i \notin \mathcal{A}_i^n$ can never be played on or off the equilibrium path. The proof is by induction on n. The $n = 1$ case is trivial. Suppose the result holds for some n. Then at any history player i's continuation payoff must lie between $\min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}_{-i}^n)} \max_{a_i\in A_i} u_i(a_i, \hat{\alpha}_{-i})$ and $\max_{a\in\mathcal{A}^n} u_i(a)$. Hence, player i will never play an action a_i for which there is no $\alpha_{-i} \in \Delta(\mathcal{A}^n)$ with

$$\begin{aligned}(1-\delta)u_i(a_i, \alpha_{-i}) &+ \delta \max_{a\in\mathcal{A}^n} u_i(a)\\ &\geq (1-\delta)\max_{\hat{a}_i\in A_i} u_i(\hat{a}_i, \alpha_{-i}) + \delta \min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}_{-i}^n)} \max_{a_i\in A_i} u_i(a_i, \hat{\alpha}_{-i}).\end{aligned}$$

This says that player i will never play an action $a_i \notin \mathcal{A}_i^{n+1}$.

We now show that if $\operatorname{int}(\bigcap_{i\in I} \bar{W}_i \cap \operatorname{co} u(\mathcal{A}^\infty)) \neq \varnothing$, then $\mathcal{A}^\infty \subseteq \operatorname{supp}(\delta)$. The argument is similar to the proof of Lemma 1.

Fix $v \in \operatorname{int}(\bigcap_{i\in I} \bar{W}_i \cap \operatorname{co} u(\mathcal{A}^\infty)) \neq \varnothing$, and let $\mu \in \Delta(\mathcal{A}^\infty)$ be such that $u(\mu) = v$ and $\mu(r) > 0$ for all $r \in \mathcal{A}^\infty$. Let α_{-i}^* be a solution to the problem $\min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}^\infty)} \max_{a_i\in A_i} u_i(\hat{a}_i, \alpha_{-i})$. Let α_{-i}^ε be the following full-support (within $\mathcal{A}^\infty$) approximation of α_{-i}^*: $\alpha_{-i}^\varepsilon = (1-\varepsilon)\alpha_{-i}^* + \varepsilon \sum_{a_{-i}\in\mathcal{A}_{-i}^\infty} \frac{a_{-i}}{|\mathcal{A}_{-i}^\infty|}$. Since $v \in \operatorname{int}(\bigcap_{i\in I} \bar{W}_i)$, there exists $\varepsilon > 0$ such that, for each $i \in I$, we have

$$v_i > \max_{a_i\in A_i} u_i\big(a_i, \alpha_{-i}^\varepsilon\big) + \frac{1-\delta}{\delta} \max_{r\in\mathcal{A}^\infty, a_i\in A_i} \big\{u_i(a_i, r_{-i}) - u_i(r)\big\}. \tag{12}$$

Fix $\varepsilon > 0$ small enough such that, for each player i, some best response to α_{-i}^ε is included in $\mathcal{A}_i^\infty$: this is always possible, as every best response to α_{-i}^* is included in $\mathcal{A}_i^\infty$.

We can now construct an equilibrium strategy μ^* with $\operatorname{supp}(\mu^*(\varnothing)) = \mathcal{A}^\infty$. The construction is similar to that in the proof of Lemma 1, with the following differences. First, μ is recommended on path, and player i's deviations are punished by recommending α_{-i}^ε to her opponents. Second, if a player deviates to an action outside $\mathcal{A}^\infty$, play reverts to an arbitrary static Nash equilibrium α^{NE} forever.

Incentive compatibility with respect to deviations within $\mathcal{A}^\infty$ follows from (12), just as it follows from (1) in the proof of Lemma 1. For incentive compatibility with respect to deviations outside $\mathcal{A}^\infty$, fix $\hat{a}_i \notin \mathcal{A}_i^\infty$ and $\alpha_{-i} \in \Delta(\mathcal{A}_{-i}^\infty)$. Note that all static Nash equilibria are contained in $\Delta(\mathcal{A}^\infty)$ and all static best responses to α_{-i} are contained in $\mathcal{A}_i^\infty$.

Therefore,

$$\begin{aligned}(1-\delta)u_i(\hat{a}_i,\alpha_{-i})+\delta u_i\big(\alpha^{\mathrm{NE}}\big) &\leq (1-\delta)u_i(\hat{a}_i,\alpha_{-i})+\delta \max_{a\in\mathcal{A}^\infty} u_i(a)\\ &< (1-\delta)\max_{\tilde{a}_i\in A_i} u_i(\tilde{a}_i,\alpha_{-i})+\delta \min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}^\infty_{-i})}\max_{a_i\in A_i} u_i(a_i,\hat{\alpha}_{-i})\\ &\quad \big(\text{since } \hat{a}_i\notin\mathcal{A}^\infty_i\big)\\ &= (1-\delta)\max_{\tilde{a}_i\in\mathcal{A}^\infty_i} u_i(\tilde{a}_i,\alpha_{-i})+\delta \min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}^\infty_{-i})}\max_{a_i\in A_i} u_i(a_i,\hat{\alpha}_{-i}).\end{aligned}$$

The last line corresponds to the deviation gain to some action $\tilde{a}_i\in\mathcal{A}^\infty_i$, which is not profitable. Hence, deviating to $\hat{a}_i\notin\mathcal{A}^\infty_i$ is not profitable either. □

Note that Proposition 3 only improves on Theorem 1 at low discount factors: for high discount factors, $A=S(A)=\mathcal{A}^\infty$, so (11) reduces to $\mathring{W}^*\neq\varnothing$. However, as $\mathring{W}^*$ can be empty only for low discount factors, this is precisely the case where an improvement is needed.[18]

To be able to use Proposition 3 to give a recursive upper bound on $E(\delta,p)$ when $\mathring{W}^*\neq\varnothing$, we must characterize $\overline{E_{\mathrm{med}}(\delta)}$ under (11). Our earlier characterization generalizes easily. In particular, the following definitions are analogous to Definitions 2 and 3.

Definition 4. For any set $V\subseteq\mathbb{R}^{|I|}$, a correlated action profile $\alpha\in\Delta(\mathrm{supp}(\delta))$ is $\mathrm{supp}(\delta)$ *enforceable* on V by a mapping $\gamma:\ \mathrm{supp}(\delta)\to V$ such that, for each player i and action $a_i\in\mathrm{supp}\,\alpha_i$,

$$\begin{aligned}&\mathbb{E}^\alpha\big[(1-\delta)u_i(a_i,a_{-i})+\delta\gamma(a_i,a_{-i})\big]\\ &\quad\geq \max_{a'_i\in A_i}\mathbb{E}^\alpha\big[(1-\delta)u_i\big(a'_i,a_{-i}\big)\big]+\delta\min_{\alpha_{-i}\in\Delta(\mathrm{supp}(\delta))}\max_{\hat{a}_i\in A_i} u_i(\hat{a}_i,\alpha_{-i}).\end{aligned}$$

Definition 5. A payoff vector $v\in\mathbb{R}^{|I|}$ is $\mathrm{supp}(\delta)$ *decomposable* on V if there exists a correlated action profile $\alpha\in\Delta(\mathrm{supp}(\delta))$ that is $\mathrm{supp}(\delta)$ enforced on V by some mapping γ such that

$$v=\mathbb{E}^\alpha\big[(1-\delta)u(a)+\delta\gamma(a)\big].$$

Let $\tilde{B}^{\mathrm{supp}(\delta)}(V)=\{v\in\mathbb{R}^{|I|}:v \text{ is } \mathrm{supp}(\delta) \text{ decomposable on } V\}$.

Let $W^{\mathrm{supp}(\delta),1}=u(\mathrm{supp}(\delta))$, let $W^{\mathrm{supp}(\delta),n}=\tilde{B}^{\mathrm{supp}(\delta)}(W^{\mathrm{supp}(\delta),n-1})$ for $n>1$, and let $W^{\mathrm{supp}(\delta),\infty}=\lim_{n\to\infty}W^{\mathrm{supp}(\delta),n}$. We have the following proposition.

Proposition 4. *If* $\mathrm{int}(\bigcap_{i\in I}\bar{W}_i\cap u(\mathcal{A}^\infty)\neq\varnothing$, *then* $\overline{E_{\mathrm{med}}(\delta)}=W^{\mathcal{A}^\infty,\infty}$.

Given that $\mathrm{supp}(\delta)=\mathcal{A}^\infty$ by Proposition 3, the proof is analogous to the proof of Theorem 2.

[18]To be clear, it is possible for W^* to be empty while $\mathcal{A}^\infty=A$. Theorem 1 and Proposition 3 only give sufficient conditions: we are not claiming that they cover every possible case.

As an example of how Propositions 3 and 4 can be applied, one can check that applying the operator S in the Bertrand example in Section 7 for any $\delta \in (\frac{1}{4}, \frac{5}{18})$ yields $\mathcal{A}^\infty = \{W, L, M\} \times \{W, L, M\}$—ruling out the efficient action profile (H, H)—and $\text{int}(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^\infty) \neq \varnothing$. We can then compute $\overline{E_{\text{med}}(\delta)}$ by applying the operator $\tilde{B}^{\mathcal{A}^\infty}$.

9.3 *What if there are more than two players?*

The condition that $\mathring{W}^* \neq \varnothing$ no longer guarantees that mediated perfect monitoring outperforms private monitoring when there are more than two players. We record this as a proposition.

Proposition 5. *There are games with $|I| > 2$ where $\mathring{W}^* \neq \varnothing$ but $\sup_{v\in\mathcal{E}(\delta)} \lambda \cdot v > \max_{v\in\overline{E_{\text{med}}(\delta)}} \lambda \cdot v$ for some nonnegative Pareto weight $\lambda \in \Lambda_+$.*

Proof. Consider the following example. There are five players and $A_i = \{a_i, b_i\}$ for $i \in \{1, 2, 3, 4\}$. Player 5 is a dummy player who takes no action and receives payoff 1 if the action profile is (a_1, b_2, a_3, a_4) or (b_1, a_2, a_3, a_4), and receives payoff 0 otherwise. The other players' payoffs are

(a_3, a_4):

	a_2	b_2
a_1	$0, 0, 0, 0$	$0, 1, -1, 1$
b_1	$1, 0, 1, -1$	$1, 1, 0, 0$

(a_3, b_4):

	a_2	b_2
a_1	$0, 0, 0, 0$	$0, 0, 0, -1$
b_1	$0, 0, 0, 1$	$0, 0, 0, 0$

(b_3, a_4):

	a_2	b_2
a_1	$0, 0, 0, 0$	$0, 0, 1, 0$
b_1	$0, 0, -1, 0$	$0, 0, 0, 0$

(b_3, b_4):

	a_2	b_2
a_1	$10, 10, 10, 10$	$10, 10, 10, 10$
b_1	$10, 10, 10, 10$	$10, 10, 10, 10$

Note that players 1 and 3 have an incentive to deviate at profile (a_1, b_2, a_3, a_4) and players 2 and 4 have an incentive to deviate at profile (b_1, a_2, a_3, a_4).

Let $\delta = \frac{\sqrt{5}-1}{2}$. Note that $\underline{u}_i = 0$ for all i; $d_1 = d_2 = 1$, $d_3 = d_4 = 10$, and $d_5 = 0$; hence, $\underline{w}_1 = \underline{w}_2 = \frac{\sqrt{5}-1}{2}$, $\underline{w}_3 = \underline{w}_4 = \frac{\sqrt{5}-1}{2}(10)$, and $\underline{w}_5 = 0$. Therefore, for example, the feasible payoff vector $(9.05, 9.05, 9, 9, 0.1)$ is an element of $\mathring{W}^*$.

Let $\lambda = (0, 0, 0, 0, 1)$: that is, we maximize player 5's payoff. We show that player 5 cannot receive payoff 1 in every period in any equilibrium with mediated perfect monitoring, while this can occur for some private monitoring structure.

We first derive the impossibility result for mediated perfect monitoring.

Claim 1. *If player 5 receives payoff 1 in every period, then the on-path continuation payoff for each player $i \in \{1, 2\}$ starting from any period is at most δ.*

Proof. If player 5 receives payoff 1 in every period, all on-path actions are either (a_1, b_2, a_3, a_4) or (b_1, a_2, a_3, a_4). Omitting player 3 and 4's actions and payoffs, let (w_1^a, w_2^a) denote continuation payoffs following (a_1, b_2), and let (w_1^b, w_2^b) denote continuation payoffs following (b_1, a_2). Then player 1's incentive compatibility constraint is $w_1^a \geq \frac{1-\delta}{\delta}$, and player 2's incentive compatibility constraint is $w_2^b \geq \frac{1-\delta}{\delta}$. Note also that $w_1^a + w_2^a = w_1^b + w_2^b = 1$ and $w_1^a, w_2^a, w_1^b, w_2^b \geq 0$.

Letting p denote the probability of (b_1, a_2, a_3, a_4), suppose we try to maximize player 1's payoff (the argument for player 2 is symmetric),

$$\max_{p, w_1^a, w_1^b} (1-\delta)p + \delta\big((1-p)w_1^a + pw_1^b\big),$$

subject to $w_1^a \geq \frac{1-\delta}{\delta}$ and $1 - w_1^b \geq \frac{1-\delta}{\delta}$. At a solution, $w_1^a = 1$ and $w_1^b = \frac{2\delta-1}{\delta}$. Hence, the objective equals

$$(1-\delta)p + \delta(1-p) + \delta p \frac{2\delta - 1}{\delta} = \delta. \qquad \triangleleft$$

Claim 2. *If player 5 receives payoff 1 in every period, then after (a_1, b_2, a_3, a_4) is played in period t, (b_1, a_2, a_3, a_4) is played with probability greater than $\frac{1}{2}$ in period $t+1$; and after (b_1, a_2, a_3, a_4) is played in period t, (a_1, b_2, a_3, a_4) is played with probability greater than $\frac{1}{2}$ in period $t+1$.*

Proof. By Claim 1, the best continuation payoff for player 1 from period $t+2$ is δ. Hence, for player 1 to play a_1 in period t, the probability p of (b_1, a_2, a_3, a_4) in period $t+1$ (conditional on (a_1, b_2, a_3, a_4) in period t), must satisfy

$$(1-\delta) \leq \delta(1-\delta)p + \delta^2 \times \delta,$$

or $p \geq \frac{1-\delta-\delta^3}{\delta(1-\delta)}$. Noting that $\delta = \frac{\sqrt{5}-1}{2}$ satisfies $1 - \delta = \delta^2$, this is equivalent to $p \geq \frac{\sqrt{5}-1}{2}$. $\triangleleft$

Claim 3. *If player 5 receives payoff 1 in every period, then the continuation payoff for each player $i \in \{3, 4\}$ starting from any period equals 0.*

Proof. As all on-path actions are either (b_1, a_2, a_3, a_4) or (a_1, b_2, a_3, a_4), player 3 and 4's continuation payoffs from any period sum to 0, and each must weakly exceed the minmax payoff of 0. Hence, both continuation payoffs must equal 0. $\triangleleft$

Claim 4. *There is no equilibrium in which player 5 receives payoff 1 in every period.*

Proof. Suppose such an equilibrium exists. Then all on-path actions are either (a_1, b_2, a_3, a_4) or (b_1, a_2, a_3, a_4). Suppose (a_1, b_2, a_3, a_4) is played in period 1. Then, by Claim 2, (b_1, a_2, a_3, a_4) is played in period 2 with probability greater than $\frac{1}{2}$. Hence, player 4 receives a negative instantaneous payoff in period 2, and by Claim 3 her continuation payoff from period 3 is nonpositive, which leaves her with a negative total continuation payoff from period 2, a contradiction. If instead (a_1, b_2, a_3, a_4) is played in period 1, then player 3 is left with a negative continuation payoff from period 2. $\triangleleft$

Turning to private monitoring, suppose that under the universal monitoring structure the mediator randomizes with equal probability between the two sequences

$$(a_1, b_2, a_3, a_4) \to (b_1, a_2, a_3, a_4) \to (a_1, b_2, a_3, a_4) \to (b_1, a_2, a_3, a_4) \to \cdots$$
$$(b_1, a_2, a_3, a_4) \to (a_1, b_2, a_3, a_4) \to (b_1, a_2, a_3, a_4) \to (a_1, b_2, a_3, a_4) \to \cdots.$$

Players 3 and 4 then always believe that players 1 and 2 play (a_1, b_2) and (b_1, a_1) with equal probability, so they are playing static best responses. Finally, with $\delta = \frac{\sqrt{5}-1}{2}$, the deviation gain of 1 for player $i \in \{1, 2\}$ is equal to the continuation payoff of $\delta + \delta^3 + \cdots = \frac{\delta}{1-\delta^2}$. □

We note that, in the proof of Proposition 5, the universal information structure can be replaced by a stationary private monitoring structure with ex ante correlation as follows: (i) let each player have an additional action c_i, with the property that all players receive payoff 0 if anyone plays c_i; (ii) specify that all players observe signal z if the action profile equals (a_1, b_2, a_3, a_4) or (b_1, a_2, a_3, a_4) and observe signal z' otherwise; (iii) construct a correlated equilibrium by having only players 1 and 2 observe the outcome of a randomizing device at the beginning of the game, and having the players play as in the proof of Proposition 5 as long as signal z realizes, switching to action profile (c_1, c_2, c_3, c_4) if signal z' realizes.

To see where the proof of Theorem 1 breaks down when $|I| > 2$, recall that the proof is based on the fact that, for any Pareto-efficient payoff v, if $v \notin W_i$ for one player i, then it must be the case that $v \in W_j$ for the other player j. This implies that incentive compatibility is a problem only for one player at a time, which lets us construct an equilibrium with perfect monitoring by basing continuation play only on that player's past recommendations (which she necessarily knows in any private monitoring structure). Alternatively, if there are more than two players, several players' incentive compatibility constraints might bind at once when we publicize past recommendations. The proof of Theorem 1 then cannot get off the ground.

We can however say some things about what happens with more than two players.

First, the argument in the proof of Proposition 3 that $\operatorname{supp}(\delta) = \mathcal{A}^\infty$ whenever $\operatorname{int}(\bigcap_{i\in I} \bar{W}_i \cap \operatorname{co} u(\mathcal{A}^\infty)) \neq \varnothing$ does not rely on $|I| = 2$. Thus, when $\operatorname{int}(\bigcap_{i\in I} \bar{W}_i \cap \operatorname{co} u(\mathcal{A}^\infty)) \neq \varnothing$, we can characterize the set of supportable actions for any number of players. This is sometimes already enough to imply a nontrivial upper bound on payoffs.

Second, Lemma 1 implies that if a payoff vector $v \in \mathring{\mathcal{F}}$ satisfies $v_i > \underline{u}_i + \frac{1-\delta}{\delta} d_i$ for all $i \in I$, then $v \in E_{\text{med}}(\delta)$. This shows that private monitoring cannot do "much" better than mediated perfect monitoring when the players are at least moderately patient (e.g., it cannot do more than order $1 - \delta$ better). It also shows that the usual full-dimensionality conditions for the perfect monitoring folk theorem (Fudenberg and Maskin 1986; Abreu et al. 1994) are not needed under mediated perfect monitoring.

Third, suppose there is a player i whose opponents $-i$ all have identical payoff functions: $\exists i : \forall j, j' \in I \setminus \{i\}$, $u_j(a) = u_{j'}(a)$ for all $a \in A$. Then the proof of Theorem 1 can be adapted to show that private monitoring cannot outperform mediated perfect monitoring in a direction where the extremal payoff vector v lies in $\bigcap_{j\in-i} W_j$ (but not necessarily in W_i). For example, if the game involves one firm and many identical consumers, then the consumers' best equilibrium payoff under mediated perfect monitoring is at least as good as under private monitoring. We can also show that the same result holds if the preferences of players $-i$ are sufficiently close to each other.

Finally, in a companion paper Sugaya and Wolitzky (2017) we investigate repeated n-player oligopoly games and show that private monitoring cannot outperform mediated perfect monitoring for any number of players and any discount factor in a class of "concave" games that includes linear Cournot and differentiated-product Bertrand competition.

10. Conclusion

This paper gives a simple sufficient condition ($\delta > \delta^*$) under which the equilibrium payoff set in a two-player repeated game with mediated perfect monitoring is a tight, recursive upper bound on the equilibrium payoff set in the same game with any imperfect private monitoring structure. There are at least three perspectives from which this result may be of interest. First, it shows that simple, recursive methods can be used to upper-bound the equilibrium payoff set in a repeated game with imperfect private monitoring at a fixed discount factor, even though the problem of recursively characterizing this set seems intractable. Second, it characterizes the set of payoffs that can arise in a repeated game for some monitoring structure. Third, it shows that information is valuable in mediated repeated games, in that players cannot benefit from imperfections in the monitoring technology when $\delta > \delta^*$.

These different perspectives on our results suggest different questions for future research. Do moderately patient players always benefit from *any* improvement in the monitoring technology, or only from going all the way to perfect monitoring? Is it possible to characterize the set of payoffs that can arise for some monitoring structure even if $\delta < \delta^*$? If we do know the monitoring structure under which the game is being played, is there a general way to use this information to tighten our upper bound? Answering these questions may improve our understanding of repeated games with private monitoring at fixed discount factors, even if a full characterization of the equilibrium payoff set in such games remains out of reach.

Appendix A: Proof of Proposition 1

Mediated perfect monitoring

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 $(1-\delta)1 = \frac{5}{6}$, her instantaneous gain from playing D rather than U. In addition, player 1 can guarantee herself a continuation payoff of 0 by always playing D, so her expected continuation payoff from playing U must exceed $\frac{1}{\delta}(\frac{5}{6}) = 5$. 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

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

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

$$\left(1-\frac{1}{6}\right)\left[\frac{3}{5}(-3)+\frac{2}{5}(3)\right]+\frac{1}{6}(3)=0.$$

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

$$\left(1-\frac{1}{6}\right)\left[\frac{3}{5}(3)+\frac{2}{5}(-3)\right]+\frac{1}{6}(0)=\frac{1}{2}.$$

Therefore, player 2 has a profitable deviation, so no such equilibrium can exist.

Private monitoring

Consider the following imperfect private monitoring structure. Player 2's action is perfectly observed. Player 1's action is perfectly observed when it equals T or B. When player 1 plays U or D, player 2 observes one of two possible private signals, m and r. Whenever player 1 plays U, player 2 observes signal m with probability 1; whenever player 1 plays D, player 2 observes signals m and r 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 $n = 0, 1, \ldots,$ player 1 plays $\frac{1}{3}U + \frac{2}{3}D$. Let $a_1(n)$ denote the realization of this mixture. In the even period $t = 2n+2$, if the previous action $a_1(n)$ equals U, then player 1 plays T; if the previous action $a_1(n)$ equals D, then player 1 plays B. If in the previous period player 1 deviated to T or B, then player 1 plays D.

Player 2's strategy. In each odd period $t = 2n+1$ with $n = 0, 1, \ldots,$ player 2 plays L. Let $y_2(n)$ denote the realization of player 2's private signal. In the even period $t = 2n+2$, if the previous private signal $y_2(n)$ equals m, then player 2 plays M; if the previous signal $y_2(n)$ equals r, then player 2 plays R. If in the previous period player 1 deviated to T or B, 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 $v = \frac{5}{6}(2) + \frac{1}{6}\frac{5}{6}(6) + \frac{1}{6^2}v$, while her payoff from D is $\frac{5}{6}(3) + \frac{1}{6}\frac{5}{6}(0) + \frac{1}{6^2}v$. Hence, player 1 is indifferent between U and D (and clearly prefers either of these 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 also independent of player 2's action. Hence, playing L is optimal for player 2.

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 in the previous period player 1 deviated to T or B, then the players play the static Nash equilibrium (D, R). 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.

Next, if player 2 observes signal $y_2(n) = m$, then his posterior belief that player 1's last action was $a_1(n) = U$ 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.

Finally, expected payoffs under this strategy profile in odd periods sum to $\frac{1}{3}(4) + \frac{2}{3}(3) = \frac{10}{3}$, and in even periods sum to 3. Therefore, per-period expected payoffs sum to

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

Three remarks on the proof: First, the various indifferences in the above argument result only because we have chosen payoffs to make the example as simple as possible. One can modify the example to make all incentives strict.[19] Second, players' payoffs are measurable with respect to their own actions and signals. In particular, the required realized payoffs for player 2 are

(Action, Signal) Pair:	(L, m)	(L, r)	(M, m)	(M, r)	(R, m)	(R, r)
Realized Payoff:	2	-2	0	0	0	0

Third, a similar argument shows that imperfect public monitoring with private strategies can also outperform mediated perfect monitoring.[20]

[19]The only nontrivial step in doing so is giving player 1 a strict incentive to mix in odd periods. This can be achieved by introducing correlation between the players' actions in odd periods.

[20]Here is a sketch: Modify the current example by adding a strategy L' for player 2, which is an exact duplicate of L as far as payoffs are concerned, but which switches the interpretation of signals m and r. Assume that player 1 cannot distinguish between L and L', and modify the equilibrium by having player 2 play $\frac{1}{2}L + \frac{1}{2}L'$ in odd periods. Then, even if the signals m and r are publicly observed, their interpretations will be private to player 2, and essentially the same argument as with private monitoring applies.

Appendix B: Proof of Lemma 6

It is useful to introduce a family of auxiliary value functions $(w^T)_{T=1}^{\infty}$ and $(w^{*,T})_{T=1}^{\infty}$, which will converge to w and w^* pointwise in r^t as $T \to \infty$. For periods $t \geq T$, define

$$w^T(r^t) = w^{\bar{\mu}}(r^t) \quad \text{and} \quad w^{*,T}(r^{t-1}) = w^{\bar{\mu}}(r^{t-1}). \tag{13}$$

For periods $t \leq T-1$, define $w^{*,T}(r^t)$, $\rho^T(r^t)$, and $w^T(r^t)$ given $w^T(r^{t+1})$ recursively, as follows. First, define

$$w^{*,T}(r^t) = (1-\delta)u(\bar{\mu}(r^t)) + \delta\mathbb{E}[w^T(r^{t+1})|r^t]. \tag{14}$$

Note that, for $t = T-1$, this definition is compatible with (13). Second, given $w^{*,T}(r^t)$, define

$$w^T(r^t) = 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}}\{\rho^T(r^t)w^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}^1\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1, \tag{15}$$

where, when $w^{\bar{\mu}}(r^t) \in W_1$, $\rho^T(r^t)$ is the largest number in $[0,1]$ such that

$$\rho^T(r^t)w_2^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}_2^1 \geq w_2^{\bar{\mu}}(r^t). \tag{16}$$

We show that $w^{*,T}$ converges to w^*.

Lemma 7. *$\lim_{T\to\infty} w^{*,T}(r^t) = w^*(r^t)$ for all $r^t \in A^t$.*

Proof. By Lemma 5, it suffices to show that $F(w^T) = w^{T+1}$. For $t \geq T+1$, (13) implies that $w^{*,T+1}(r^{t-1}) = w^{\bar{\mu}}(r^{t-1})$. Given w^T, $w^*(w_2^T)$ is the value calculated according to (6). Since $w^T(r^t) = w^{\bar{\mu}}(r^t)$ by (13), we have $w^*(w_2^T)(r^{t-1}) = w^{\bar{\mu}}(r^{t-1})$ by (6). Hence,

$$w^*(w_2^T)(r^{t-1}) = w^{*,T+1}(r^{t-1}). \tag{17}$$

For $t \leq T$, by (14), we have

$$w^{*,T+1}(r^t) = (1-\delta)u(\bar{\mu}(r^t)) + \delta\mathbb{E}[w^{T+1}(r^{t+1})|r^t].$$

By (15),

$$\begin{aligned} w^{T+1}(r^t) &= 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}}\{\rho^{T+1}(r^t)w^{*,T+1}(r^t) + (1-\rho^{T+1}(r^t))\bar{w}^1\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1 \\ &= 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}}\{\rho^{T+1}(r^t)w^*(w^T)(r^t) + (1-\rho^{T+1}(r^t))\bar{w}^1\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1, \end{aligned}$$

where the second equality follows from (17) for $t = T$, and follows by induction for $t < T$.

Recall that ρ^{T+1} is defined by (16), while $\rho(w^T)(r^t)$ is defined in (8) using $w^* = w^*(w^T)$. Since $w^{*,T+1}(r^t) = w^*(w_2^T)(r^t)$, we have $\rho^{T+1}(r^t) = \rho(w^T)(r^t)$. Hence,

$$\begin{aligned} w^{T+1}(r^t) &= 1_{\{w^{\bar{\mu}}(r^t)\in W_1\}}\{\rho(w^T)(r^t)w^*(w^T)(r^t) + (1-\rho(w^T)(r^t))\bar{w}^1\} + 1_{\{w^{\bar{\mu}}(r^t)\notin W_1\}}\bar{w}^1. \\ &= F(w^T)(r^t), \end{aligned}$$

as desired. □

As $w^{*,T}(r^t)$, $w^T(r^t)$, and $\rho^T(r^t)$ converge to $w^*(r^t)$, $w(r^t)$, and $\rho(r^t)$ by Lemma 7, the following lemma implies Lemma 6.

Lemma 8. *For all $t = 1, \dots, T-1$, if $w^{\bar{\mu}}(r^t) \in W_1$, then $\rho^T(r^t)w^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}^1$ Pareto dominates $w^{\bar{\mu}}(r^t)$.*

Proof. For $t = T-1$, the claim is immediate since $w^{*,T}(r^t) = w^{\bar{\mu}}(r^t)$ and so $\rho^T(r^t) = 1$.

Suppose that the claim holds for each period $\tau \geq t+1$. We show that it also holds for period t. By construction, $\rho^T(r^t)w_2^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}_2^1 \geq w_2^{\bar{\mu}}(r^t)$. Thus, it suffices to show that $\rho^T(r^t)w_1^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}_1^1 \geq w_1^{\bar{\mu}}(r^t)$.

Note that

$$
\begin{aligned}
&w^{*,T}(r^t)\\
&\quad = (1-\delta)u(\bar{\mu}(r^t)) + \delta\mathbb{E}[w^T(r^{t+1})|r^t]\\
&\quad = (1-\delta)u(\bar{\mu}(r^t))\\
&\qquad + \delta\left\{\begin{array}{l}\sum_{r^{t+1}: w^{\bar{\mu}}(r^{t+1})\in W_1}\Pr^{\bar{\mu}}(r^{t+1}|r^t)\{\rho^T(r^{t+1})w^{*,T}(r^{t+1}) + (1-\rho^T(r^{t+1}))\bar{w}^1\}\\ \quad + \sum_{r^{t+1}:\ w^{\bar{\mu}}(r^{t+1})\notin W_1}\Pr^{\bar{\mu}}(r^{t+1}|r^t)\bar{w}^1\end{array}\right\},
\end{aligned}
$$

while

$$
w^{\bar{\mu}}(r^t) = (1-\delta)u(\bar{\mu}(r^t)) + \delta\sum_{r^{t+1}}\Pr^{\bar{\mu}}(r^{t+1}|r^t)w^{\bar{\mu}}(r^t).
$$

Hence,

$$
\begin{aligned}
&w^{*,T}(r^t) - w^{\bar{\mu}}(r^t)\\
&\quad = \delta\left\{\begin{array}{l}\sum_{r^{t+1}:\ w^{\bar{\mu}}(r^{t+1})\in W_1}\Pr^{\bar{\mu}}(r^{t+1}|r^t)\\ \quad \times\{\rho^T(r^{t+1})w^{*,T}(r^{t+1}) + (1-\rho^T(r^{t+1}))\bar{w}^1 - w^{\bar{\mu}}(r^{t+1})\}\\ \quad + \sum_{r^{t+1}:\ w^{\bar{\mu}}(r^{t+1})\notin W_1}\Pr^{\bar{\mu}}(r^{t+1}|r^t)\{\bar{w}^1 - w^{\bar{\mu}}(r^{t+1})\}\end{array}\right\}.
\end{aligned}
$$

When $w^{\bar{\mu}}(r^{t+1}) \in W_1$, the inductive hypothesis implies that

$$
\rho^T(r^{t+1})w^{*,T}(r^{t+1}) + (1-\rho^T(r^{t+1}))\bar{w}^1 - w^{\bar{\mu}}(r^{t+1}) \geq 0.
$$

In addition, note that

$$
\sum_{r^{t+1}:\ w^{\bar{\mu}}(r^{t+1})\notin W_1}\Pr^{\bar{\mu}}(r^{t+1}|r^t)\{\bar{w}^1 - w^{\bar{\mu}}(r^{t+1})\} = l(r^t)(\bar{w}^1 - \tilde{w}(r^t))
$$

for some number $l(r^t) \geq 0$ and vector $\tilde{w}(r^t) \notin W_1$. In total, we have

$$
w^{*,T}(r^t) = w^{\bar{\mu}}(r^t) + l(r^t)(\bar{w}^1 - \hat{w}(r^t)) \tag{18}
$$

for some number $l(r^t) \geq 0$ and vector $\hat{w}(r^t) \leq \tilde{w}(r^t) \notin W_1$. Since $\bar{w}_1^1 \geq \hat{w}_1(r^t)$, if $\bar{w}_1^1 \geq w_1^{\bar{\mu}}(r^t)$, then (18) implies that $\min\{w_1^{*,T}(r^t), \bar{w}_1^1\} \geq w_1^{\bar{\mu}}(r^t)$ and, therefore, $\rho^T(r^t)w_1^{*,T}(r^t) +$

$(1-\rho^T(r^t))\bar{w}_1^1 \geq w_1^{\bar{\mu}}(r^t)$. In addition, if $w^{\bar{\mu}}(r^{t+1}) \in W_1$ with probability 1, then the inductive hypothesis implies that $w_2^{*,T}(r^t) \geq w_2^{\bar{\mu}}(r^t)$, and therefore $\rho^T(r^t) = 1$ and

$$\begin{aligned}\rho^T(r^t)w_1^{*,T}(r^t) + (1-\rho^T(r^t))\bar{w}_1^1 &= w_1^{*,T}(r^t) \\ &= w_1^{\bar{\mu}}(r^t) + l(r^t)(\bar{w}_1^1 - \hat{w}_1(r^t)) \\ &\geq w_1^{\bar{\mu}}(r^t).\end{aligned}$$

Hence, it remains only to consider the case where $\bar{w}_1^1 < w_1^{\bar{\mu}}(r^t)$ and $l(r^t) > 0$.

In this case, take a normal vector λ^1 of the supporting hyperplane of $\mathcal{F}$ at $\bar{w}^1$. We have $\lambda_1^1 \geq 0$ and $\lambda_2^1 > 0$, and in addition (as $\hat{w}(r^t) \leq \tilde{w}(r^t) \in \mathcal{F}$ and $w^{\bar{\mu}}(r^t) \in \mathcal{F}$),

$$\begin{aligned}\lambda^1 \cdot (\bar{w}^1 - \hat{w}(r^t)) &\geq 0, \\ \lambda^1 \cdot (\bar{w}^1 - w^{\bar{\mu}}(r^t)) &\geq 0.\end{aligned}$$

As $\bar{w}_1^1 - \hat{w}_1(r^t) > 0$ and $\bar{w}_1^1 - w_1^{\bar{\mu}}(r^t) < 0$, we have

$$\frac{\hat{w}_2(r^t) - \bar{w}_2^1}{\bar{w}_1^1 - \hat{w}_1(r^t)} \leq \frac{\lambda_1^1}{\lambda_2^1} \leq \frac{\bar{w}_2^1 - w_2^{\bar{\mu}}(r^t)}{w_1^{\bar{\mu}}(r^t) - \bar{w}_1^1}.$$

Next, by (18), the slope of the line from $w^{\bar{\mu}}(r^t)$ to $w^{*,T}(r^t)$ equals the slope of the line from $\hat{w}(r^t)$ to $\bar{w}^1$. Hence,

$$\frac{w_2^{\bar{\mu}}(r^t) - w_2^{*,T}(r^t)}{w_1^{*,T}(r^t) - w_1^{\bar{\mu}}(r^t)} \leq \frac{\bar{w}_2^1 - w_2^{\bar{\mu}}(r^t)}{w_1^{\bar{\mu}}(r^t) - \bar{w}_1^1}.$$

In this inequality, the denominator of the left-hand side and the numerator of the right-hand side are both positive, $w_1^{*,T}(r^t) > w_1^{\bar{\mu}}(r^t)$ by (18) and $l(r^t) > 0$, while $\bar{w}_2^1 > w_2^{\bar{\mu}}(r^t)$ because $w^{\bar{\mu}}(r^t) \in W_1$ and $w^{\bar{\mu}}(r^t) \neq \bar{w}_2^1$. Therefore, the inequality is equivalent to

$$\frac{w_2^{\bar{\mu}}(r^t) - w_2^{*,T}(r^t)}{\bar{w}_2^1 - w_2^{\bar{\mu}}(r^t)} \leq \frac{w_1^{*,T}(r^t) - w_1^{\bar{\mu}}(r^t)}{w_1^{\bar{\mu}}(r^t) - \bar{w}_1^1}.$$

Now, let $q \in [0, 1]$ be the number such that

$$q w_1^{*,T}(r^t) + (1-q)\bar{w}_1^1 = w_1^{\bar{\mu}}(r^t).$$

Note that

$$1 - \rho^T(r^t) = \frac{w_2^{\bar{\mu}}(r^t) - w_2^{*,T}(r^t)}{\bar{w}_2^1 - w_2^{\bar{\mu}}(r^t)},$$

while

$$1 - q = \frac{w_1^{*,T}(r^t) - w_1^{\bar{\mu}}(r^t)}{w_1^{\bar{\mu}}(r^t) - \bar{w}_1^1}.$$

Hence, $\rho^T(r^t) > q$. Finally, we have seen that $\bar{w}_1^1 \leq w_1^{\bar{\mu}}(r^t) \leq w_1^{*,T}(r^t)$, so we have

$$\rho^T(r^t) w_1^{*,T}(r^t) + (1 - \rho^T(r^t))\bar{w}_1^1 \geq q w_1^{*,T}(r^t) + (1-q)\bar{w}_1^1 = w_1^{\bar{\mu}}(r^t). \qquad \square$$

References

Abreu, D., P. K. Dutta, and L. Smith (1994), "The folk theorem for repeated games: A NEU condition." *Econometrica*, 62, 939–948. [720]

Abreu, D., D. Pearce, and E. Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [691]

Amarante, M. (2003), "Recursive structure and equilibria in games with private monitoring." *Economic Theory*, 22, 353–374. [694]

Athey, S. and K. Bagwell (2001), "Optimal collusion with private information." *RAND Journal of Economics*, 32, 428–465. [691]

Awaya, Y. and V. Krishna (2015), "On communication and collusion." *American Economic Review*, 106, 285–315. [694]

Ben-Porath, E. and M. Kahneman (1996), "Communication in repeated games with private monitoring." *Journal of Economic Theory*, 70, 281–297. [694]

Bergemann, D. and S. Morris (2013), "Robust predictions in games with incomplete information." *Econometrica*, 81, 1251–1308. [694]

Compte, O. (1998), "Communication in repeated games with imperfect private monitoring." *Econometrica*, 66, 597–626. [694]

Dhillon, A. and J.-F. Mertens (1996), "Perfect correlated equilibria." *Journal of Economic Theory*, 68, 279–302. [696]

Forges, F. (1986), "An approach to communication equilibria." *Econometrica*, 54, 1375–1385. [696]

Forges, F. (2009), "Correlated Equilibrium and Communication in Games." In *Encyclopedia of Complexity and Systems Science* (R. Meyers, ed.), 1587–1596, Springer, New York. [714]

Fuchs, W. (2007), "Contracting with repeated moral hazard and private evaluations." *American Economic Review*, 97, 1432–1448. [691]

Fudenberg, D. and D. Levine (1983), "Subgame-perfect equilibria of finite- and infinite-horizon games." *Journal of Economic Theory*, 31, 251–268. [701]

Fudenberg, D., D. Levine, and E. Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [691]

Fudenberg, D. and E. Maskin (1986), "The folk theorem in repeated games with discounting or with incomplete information." *Econometrica*, 54, 533–554. [720]

Gossner, O. (2000), "Comparison of information structures." *Games and Economic Behavior*, 30, 44–63. [694]

Green, E. J. and R. H. Porter (1984), "Noncooperative collusion under imperfect price information." *Econometrica*, 52, 87–100. [691]

Heller, Y., E. Solan, and T. Tomala (2012), "Communication, correlation and cheap-talk in games with public information." *Games and Economic Behavior*, 74, 222–234. [714]

Kandori, M. (1991), "Cooperation in finitely repeated games with imperfect private information." Report. [698]

Kandori, M. (2002), "Introduction to repeated games with private monitoring." *Journal of Economic Theory*, 102, 1–15. [692]

Kandori, M. and H. Matsushima (1998), "Private observation, communication and collusion." *Econometrica*, 66, 627–652. [694]

Kandori, M. and I. Obara (2010), "Towards a belief-based theory of repeated games with private monitoring: An application of POMDP." Report. [694]

Lehrer, E. (1992), "Correlated equilibria in two-player repeated games with nonobservable actions." *Mathematics of Operations Research*, 17, 175–199. [693, 694]

Lehrer, E., D. Rosenberg, and E. Shmaya (2010), "Signaling and mediation in games with common interests." *Games and Economic Behavior*, 68, 670–682. [694]

Levin, J. (2003), "Relational incentive contracts." *American Economic Review*, 93, 835–857. [691]

MacLeod, B. W. (2003), "Optimal contracting with subjective evaluation." *American Economic Review*, 93, 216–240. [691]

Mailath, G. J., S. A. Matthews, and T. Sekiguchi (2002), "Private strategies in finitely repeated games with imperfect public monitoring." *The B.E. Journal of Theoretical Economics*, 2, 1–23. [698]

Mailath, G. J. and L. Samuelson (2006), *Repeated Games and Reputations*. Oxford University Press, Oxford. [698, 704]

Mertens, J.-F., S. Sorin, and S. Zamir (2015), *Repeated Games*. Cambridge University Press, Cambridge. [693, 694]

Miyahara, Y. and T. Sekiguchi (2013), "Finitely repeated games with monitoring options." *Journal of Economic Theory*, 148, 1929–1952. [698]

Myerson, R. B. (1986), "Multistage games with communication." *Econometrica*, 54, 323–358. [696]

Pai, M. M., A. Roth, and J. Ullman (2014), "An anti-folk theorem for large repeated games with imperfect monitoring." Report. [694]

Phelan, C. and A. Skrzypacz (2012), "Beliefs and private monitoring." *Review of Economic Studies*, 79, 1637–1660. [694]

Phelan, C. and E. Stacchetti (2001), "Sequential equilibria in a Ramsey tax model." *Econometrica*, 69, 1491–1518. [691]

Rahman, D. (2012), "But who will monitor the monitor?" *American Economic Review*, 102, 2767–2797. [693]

Rahman, D. (2014), "The power of communication." *American Economic Review*, 104, 3737–3751. [713]

Rahman, D. and I. Obara (2010), "Mediated partnerships." *Econometrica*, 78, 285–308. [693]

Renault, J. and T. Tomala (2004), "Communication equilibrium payoffs in repeated games with imperfect monitoring." *Games and Economic Behavior*, 49, 313–344. [693, 694]

Sekiguchi, T. (2002), "Existence of nontrivial equilibria in repeated games with imperfect private monitoring." *Games and Economic Behavior*, 40, 299–321. [698]

Stigler, G. J. (1964), "A theory of oligopoly." *Journal of Political Economy*, 72, 44–61. [691]

Sugaya, T. and A. Wolitzky (2017), "Bounding payoffs in repeated games with private monitoring: n-player games." Report. [698, 721]

Tomala, T. (2009), "Perfect communication equilibria in repeated games with imperfect monitoring." *Games and Economic Behavior*, 67, 682–694. [694]

Co-editor George J. Mailath handled this manuscript.

Manuscript received 26 August, 2015; final version accepted 25 April, 2016; available online 25 April, 2016.