On the Equilibrium Payoff Set in Repeated Games with Imperfect Private Monitoring*

Takuo Sugaya and Alexander Wolitzky
Stanford Graduate School of Business and MIT

December 10, 2014

Abstract

We provide simple sufficient conditions for the existence of a tight, recursive upper bound on 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 conditions, this set describes precisely those payoff vectors that arise in equilibrium for some private monitoring structure.

PRELIMINARY. COMMENTS WELCOME.

1 Introduction

Like many dynamic economic models, repeated games are typically studied using recursive methods. In an incisive paper, Abreu, Pearce, and Stacchetti (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, Levine, and Maskin (1994) and others) led to fresh perspectives on problems like collusion (Green and Porter, 1984; Athey and Bagwell, 2001), relational

*This replaces our earlier paper, "Perfect Versus Imperfect Monitoring in Repeated Games." For helpful comments, we thank Gabriel Carroll, Michihiro Kandori, Hitoshi Matsushima, Stephen Morris, Tadashi Sekiguchi, and Andrzej Skrzypacz.

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 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 simple conditions for the existence of a tight, recursive upper bound on this set. Equivalently, under these conditions we give a recursive characterization of the set of payoffs that can be attained in equilibrium for *some* private monitoring structure.[1] Thus, from the perspective of an observer who knows the monitoring structure, our results give an upper bound on how well players can do in a repeated game; while from the perspective of an observer who does not know the monitoring structure, our results exactly characterize how well the players can do.

Throughout the paper, 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 (or, more precisely, the *closure* of the *strict* equilibrium payoff set with this information structure). We thus show that the equilibrium payoff set with private monitoring has a simple, recursive upper bound by establishing two main results:

1. Under some conditions, the equilibrium payoff set with perfect monitoring and a mediator is indeed an upper bound on the equilibrium payoff set with any private monitoring structure.

2. The equilibrium payoff set with perfect monitoring and a mediator has a simple recur-

[1] The required private monitoring structure may be *stochastic*, meaning that the distribution of signals can depend on everything that has happened in the past, rather than only on current actions. With a repeated (non-stochastic) monitoring structure, the same characterization obtains if the players can communicate through cheap talk. See Section 7.2.

sive structure.

At first glance, 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 imperfect private monitoring can sometimes outperform (mediated) perfect monitoring. Our first result provides sufficient conditions for this not to happen. Thus, another contribution of our paper is pointing out that perfect monitoring is not necessarily the optimal monitoring structure in a repeated game (even if it is advantaged by giving players access to a mediator), while also giving sufficient conditions under which perfect monitoring is indeed optimal.

Our sufficient condition for mediated perfect monitoring to outperform any private monitoring structure is that there is a feasible continuation 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 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. (They need not be extremely patient: none of our main results concern the limit $\delta \to 1$.) The reason why this condition is sufficient for mediated perfect monitoring to outperform private monitoring in two-player games is fairly subtle, so we postpone a detailed discussion.

Our second main result also involves some subtleties. 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. On the other hand, *with* a mediator—who makes private action recommendations to the players—private strategies play a crucial role, and APS's characterization does not apply. We nonetheless show that the sequential equilibrium payoff set with perfect monitoring and a mediator does have a simple recursive structure. Under the sufficient conditions for our first result, a recursive characterization is obtained by replacing APS's generating operator B with what we call a *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 rewards are drawn from W and deviators are minmaxed.[2] To see intuitively 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 $\varepsilon / |A|$; and 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.[3] (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.) Our recursive characterization of the equilibrium payoff set with perfect monitoring and a mediator takes into account the possibility of minmaxing deviators in this way.

Our paper is certainly not the first to develop recursive methods for private monitoring repeated games. In an early and influential paper on repeated private monitoring games, Kandori and Matsushima (1998) augment repeated private monitoring games with opportunities for public communication among players, and provide a recursive characterization of the equilibrium payoff set for a subclass of equilibria 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. However, neither paper provides a recursive upper bound on the entire sequential equilibrium payoff set at a fixed discount factor in the repeated game.[4] Amarante (2003) does give a recursive characterization of

[2] The equilibrium payoff set with perfect monitoring and a mediator has a recursive structure whether or not the sufficient conditions for our first result are satisfied, but the characterization is somewhat more complicated in the general case. See Section 7.1.

[3] In this construction, the mediator *virutally implements* the target action profile. For other applications of virtual implementation in games with a mediator, see Rahman and Obara (2010) and Rahman (2012).

[4] Ben-Porath and Kahneman (1996) and Compte (1998) also prove folk theorems for repeated private monitoring games with communication, but they do not emphasize recursive methods away from the $\delta \to 1$ limit.

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. In contrast, our upper bound is recursive in payoff space, just like in APS.

Recently, 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. Their results do not give a recursive characterization or upper bound on the equilibrium payoff set. The type of recursive methods used in their papers is also different: their methods for checking whether a given strategy profile is an equilibrium involve a recursion on the sets of beliefs that players can have about each other's states, rather than a recursion on payoffs.

Like our paper, a working paper by Cherry and Smith (2011) tries to give a recursive upper bound on the equilibrium payoff set in private monitoring repeated games by comparing this set with the equilibrium payoff set with perfect monitoring and mediation or extensive-form correlation. However, to the best of our knowledge, Cherry and Smith do not provide conditions under which the equilibrium payoff set with perfect monitoring and a mediator is actually an upper bound on the equilibrium payoff set with private monitoring (as we do in Theorem 1), and they also do not give a recursive characterization of the sequential equilibrium payoff set with perfect monitoring and a mediator that takes private strategies into account (as we do in Theorem 2). Thus, while our motivation is very similar to Cherry and Smith's, our results are quite different from theirs. Another recent paper by Awaya and Krishna (2014) derives an upper bound on payoffs in a repeated Bertrand game for a particular private monitoring structure. They do not address the question of bounding equilibrium payoffs for general games, or that of bounding equilibrium payoffs independently of the monitoring structure.

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 in a repeated game 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.

The rest of the paper is organized as follows. Section 2 describes our models of repeated games with imperfect private monitoring and repeated games with perfect monitoring and a mediator, which are standard. Section 3 gives an example showing that private monitoring can sometimes outperform perfect monitoring with a mediator. Section 4 presents our first main result, which gives sufficient conditions for such examples not to exist. Section 5 presents our second main result: a simple recursive characterization of the equilibrium payoff set with perfect monitoring and a mediator. Combining the results of Sections 4 and 5 gives the desired recursive upper bound on the equilibrium payoff set with private monitoring. Section 6 illustrates the calculation of the upper bound with an example. Finally, Section 7 discusses partial versions of our results that apply when our sufficient conditions do not hold, as in the case of more than two players, as well as the tightness of our upper bound.

2 Model

A stage game $G = \left(I, (A_i)_{i \in I}, (u_i)_{i \in I}\right)$ is repeated in periods $t = 1, 2, \ldots$, where $I = \{1, \ldots, |I|\}$ is the set of players, A_i is the finite set of player i's actions, and $u_i : A \to \mathbb{R}$ is player i's payoff function. Players maximize expected discounted payoffs with common discount factor δ. We compare the equilibrium payoff sets in this repeated game under *private monitoring* and *mediated perfect monitoring*.

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 $z_t = (z_{i,t})_{i \in I} \in (Z_i)_{i \in I} = Z$ 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 of player i's 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 $\emptyset$. A (behavior) strategy of player i's is a map $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(A_i)$.

We do not impose the common assumption that a player's payoff is measurable with respect to her own action and signal (i.e., that players observe their own payoffs), because none of our results need this assumption. All examples considered in the paper do however satisfy this measurability condition.

The solution concept is sequential equilibrium. More precisely, a belief system of player i's is a map $\beta_i : \bigcup_{t=1}^{\infty} H_i^t \to \bigcup_{t=1}^{\infty} \Delta(H^t)$ satisfying $\operatorname{supp} \beta_i(h_i^t) \subseteq (h_i^t, H_{-i}^t)$ for all t; we also write $\beta_i(h^t | h_i^t)$ for the probability of h^t under $\beta_i(h_i^t)$. We say that an assessment (σ, β) constitutes a sequential equilibrium if the following two conditions are satisfied:

1. [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 $\beta_i(h_i^t)$.

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

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

   (b) Conditional probabilities converge to β (pointwise in t): For all $\varepsilon > 0$ and t, there exists N such that, for all $n > N$,
$$\left| \frac{\Pr^{\sigma^n}(h_i^t, h_{-i}^t)}{\sum_{\tilde{h}_{-i}^t} \Pr^{\sigma^n}(h_i^t, \tilde{h}_{-i}^t)} - \beta_i(h_i^t, h_{-i}^t \mid h_i^t) \right| < \varepsilon \text{ for all } i \in I, h_i^t \in H_i^t, h_{-i}^t \in H_{-i}^t.$$

We choose this relatively permissive definition of consistency (requiring that strategies and beliefs converge only pointwise in t) to make our results upper-bounding the sequential equilibrium payoff set stronger. The results with a more restrictive definition (requiring uniform convergence) would be essentially the same.

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 of the mediator's 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 of the mediator's is a map $\mu : \bigcup_{t=1}^{\infty} H_m^t \to \Delta(M)$. A period t history of player i's is an element of $H_i^t = (M_i \times A)^{t-1} \times M_i$, with typical element $h_i^t = \left((m_{i,\tau}, a_\tau)_{\tau=1}^{t-1}, m_{i,t}\right)$, where $H_i^1 = M_i$. A strategy of player i's 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. 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 a "stochastic private monitoring structure" where $Z_i = M_i \times A$ and $p(\cdot | h_m^{t+1})$ coincides with perfect monitoring of actions with messages given by $\mu(h_m^{t+1})$.

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

1. [Messages are action recommendations] $M = A$.

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

Without loss of generality, we restrict attention to such *obedient* equilibria throughout.

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.

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 be a (recursive) upper bound on the equilibrium payoff set with private monitoring. Before giving these main results, we 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	$0, -3$	$0, 3$	$0, 3$

We show the following:

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

Proof. See appendix. ■

In the constructed private monitoring structure in the proof of Proposition 1, players' payoffs are measurable with respect to their own actions and signals. In addition, a similar argument shows that imperfect public monitoring (with private strategies) can also outperform mediated perfect monitoring. This shows that some conditions are also required to guarantee that the sequential equilibrium payoff set with mediated perfect monitoring is an upper bound on the sequential equilibrium payoff set with imperfect public monitoring. Since imperfect public monitoring is a special case of imperfect private monitoring, our sufficient conditions for private monitoring are enough.

The intuition for Proposition 1 is that 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 probability. With perfect monitoring, player 2 (column player, "he") always "sees (T, M) coming" after (U, L), and will therefore deviate to L. With private monitoring, player 2 may not know whether (U, L) has just occurred, and therefore may be unsure of whether the next action profile will be (T, M) or (B, M), which gives him the necessary incentive to play M. More generally, 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.[5]

To preview our results, we will show that private monitoring cannot outperform mediated perfect monitoring when there exists a feasible payoff vector that is appealing enough to both players that neither is tempted to deviate when it is promised to them in continuation if they conform. This condition is violated in the current example because, for example, no feasible continuation payoff for player 2 is high enough to induce him to respond to T with M rather than L.

[5] As far as we know, the observation that players can benefit from imperfections in monitoring even in the presence of a mediator is original. Examples by Kandori (1991), Sekiguchi (2002), Mailath, Matthews, and Sekiguchi (2002), and Miyahara and Sekiguchi (2013) show that players can benefit from imperfect monitoring in finitely repeated games. 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 an old one.

4 A Sufficient Condition for $\overline{E_{\text{med}}(\delta)}$ to Give an Upper Bound

This section provides our sufficient condition for the (closure of the strict) equilibrium payoff set with mediated perfect monitoring to upper-bound the equilibrium payoff set with private monitoring. The condition applies only for two-player games, and it implies that all action profiles can be played (with some probability) in equilibrium. In Section 7, we discuss what happens when these conditions are relaxed.

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 the minmax problem for player i. Let d_i be player i's greatest possible gain from a deviation at any recommendation profile, so that

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

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

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

Finally, let

$$W_i = \left\{ w \in \mathbb{R}^N : w_i \geq \underline{w}_i \right\}.$$

Let $u(A)$ be the convex hull of the set of feasible payoffs; let $E(\delta, p)$ be the set of (possibly weak) equilibrium payoffs with private monitoring structure p; and recall that $E_{\text{med}}(\delta)$ is the set of strict equilibrium payoffs with perfect monitoring and a mediator. We show that, if there are two players and $\bigcap_{i \in I} W_i \cap u(A)$ has a non-empty interior (as a subspace of $u(A)$), then $\overline{E_{\text{med}}(\delta)}$ is an upper bound for the Pareto frontier of $E(\delta, p)$, for any private monitoring structure p. For notational convenience, denote the interior of $\bigcap_{i\in I} W_i \cap u(A)$ as a subspace of $u(A)$ by

$$\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right).$$

The following is our first main result. Note that $E(\delta, p)$ is closed, as we use the product topology on assessments (Fudenberg and Levine, 1983), so the maxima in the theorem exist.

Theorem 1 *If $|I| = 2$ and $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$, then for every private monitoring structure p and every non-negative Pareto weight $\lambda \in \Lambda_+ \equiv \{\lambda \in \mathbb{R}^2_+ : \|\lambda\| = 1\}$, we have*

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

Note that the condition $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ is a joint restriction on the stage game and the discount factor. In particular, as δ increases to one, $\underline{w}_i$ decreases to $\underline{u}_i$, so $\bigcap_{i\in I} W_i \cap u(A)$ converges to the feasible and individually rational set. Hence, $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right)$ is non-empty for high enough δ whenever there exists a feasible and strictly individually rational payoff vector. Of course, $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right)$ will typically also be non-empty for discount factors much less than one. Theorem 1 says that, whenever this is the case, the equilibrium payoff set with mediated perfect monitoring upper-bounds the equilibrium payoff set with private monitoring in two-player games.

The idea of the proof is as follows. Let $\mathcal{E}(\delta)$ be the equilibrium payoff set in the mediated repeated game with the following "universal" information 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 $a_{i,t}$. With this monitoring structure, the mediator can clearly replicate any private monitoring structure p by setting $\mu(h^t_m)$ equal to $p(\cdot|a_{t-1})$ for every history $h^t_m = (r_\tau, a_\tau)^{t-1}_{\tau=1}$. It particular, we have $E(\delta, p) \subseteq \mathcal{E}(\delta)$ for every p,[6] so to prove Theorem 1 it suffices to show that

$$\max_{v \in \mathcal{E}(\delta)} \lambda \cdot v \leq \max_{v \in \overline{E_{\text{med}}(\delta)}} \lambda \cdot v. \tag{1}$$

(The maximum over $\mathcal{E}(\delta)$ is well defined since $\mathcal{E}(\delta)$ is closed, by the same reasoning as for $E(\delta, p)$.)

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.[7] To see why this is possible (when our sufficient conditions are satisfied), 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. (This is assuming that deviators are always minmaxed, which, as we will see, is always possible if $\text{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$.) Next, if $\text{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$ and $|I| = 2$, then if a Pareto-efficient payoff vector v lies outside of W_i for one player (say player 2), it must then lie inside of 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 $r^t = (r_\tau)_{\tau=1}^{t-1}$, but that the mediator then issues period t recommendations according to the original equilibrium distribution of recommendations conditional on *player 2's past recommendations* $r_2^t = (r_{2,\tau})_{\tau=1}^{t-1}$ *only.* As in h_m^t, the superscript t in r^t means that r^t is the information available at the beginning of period t. We claim that doing this violates neither player's obedience constraint: Player 1's obedience constraint is easy to satisfy, as 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

[6] In light of this fact, the reader may wonder why we do not simply take $\mathcal{E}(\delta)$ rather than $\overline{E_{\text{med}}(\delta)}$ to be our bound on equilibrium payoffs with private monitoring. The answer is that, as far as we know, $\mathcal{E}(\delta)$ does not admit a recursive characterization, while we recursively characterize $\overline{E_{\text{med}}(\delta)}$ in Section 5.

[7] 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.

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 clear how we can 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, but promise it with higher probability. Finally, since the feasible payoff set is convex, this "compromise" continuation payoff vector is also acceptable to player 1.

The proof is given over the next two subsections. Section 4.1 gives preliminary results, which are also used in Section 5. In particular, it presents the key result that $\overline{E_{\text{med}}(\delta)}$ contains any payoff vector that can be enforced by the threat of minmaxing deviators (Lemma 2). Section 4.2 gives the main construction. The proof of another crucial Lemma—Lemma 4—is deferred to the appendix.

4.1 Proof of Theorem 1: Preliminaries

We start by introducing the notion of a "full support equilibrium": 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 exists $(r_{-i,\tau})_{\tau=1}^{t-1}$ with $\Pr^\mu(r_\tau | (r_{\tau'}, a_{\tau'})_{\tau'=1}^{\tau-1}) > 0$ for each $\tau = 1, ..., t-1$, there exists $(\bar{r}_{-i,\tau})_{\tau=1}^{t-1}$ such that for each $\tau = 1, ..., t-1$, we have

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

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). Note that, if the equilibrium has full support, player i never believes that any of the other players has deviated.

The following lemma says that all payoffs in $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right)$ are attainable in a strict full-support equilibrium with mediated perfect monitoring.

Lemma 1 *For all $v \in \operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right)$, there exists a strict full-support equilibrium with mediated perfect monitoring with payoff v. In particular,* $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \subseteq E_{\text{med}}(\delta)$.

Proof. For each $v \in \operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right)$, there exists $\mu \in \Delta(A)$ such that $u(\mu) = v$ and $\mu(r) > 0$ for all $r \in A$. On the other hand, for each $i \in I$ and $\varepsilon \in (0,1)$, consider the following full-support approximation of the minmax strategy α_{-i}^*: $\alpha_{-i}^\varepsilon = (1-\varepsilon)\,\alpha_{-i}^* + \varepsilon \sum_{a_{-i}\in A_{-i}} \frac{a_{-i}}{|A_{-i}|}$. Since $v \in \operatorname{int}\left(\bigcap_{i\in I} W_i\right)$, there exists $\varepsilon \in (0,1)$ such that, for each $i \in I$, we have

$$v_i > \max_{a_i\in A_i} u_i(a_i, \alpha_{-i}^\varepsilon) + \frac{1-\delta}{\delta} d_i. \tag{2}$$

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. Intuitively, J is the set of players who have deviated so far.

If the state J is equal to $\emptyset$ (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 α_{-i}^ε (a full-support approximation of the minmax punishment for the deviator), and recommends some best response to α_{-i}^ε to player i. Finally, if $|J| \geq 2$ (several players have deviated), then for each $i \in J$, the mediator recommends some best response to α_{-i}^ε, while she recommends r_{-J} to the other players $-J$ according to $\frac{a_{-J}}{|A_{-J}|}$ (a random, full support recommendation). The state transits as follows: if the current state is J and players J' deviate, then the state transits to $J \cup J'$.[8]

[8] We thank Gabriel Carroll for suggestions which helped simplify this construction.

Player i's strategy is to follow the recommendation $r_{i,t}$ in period t. She believes that the mediator's state is $\emptyset$ 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. 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, (2) implies that player i has a strict incentive to follow her recommendation r_i at any recommendation profile $r \in A$. Hence, she has a strict incentive to follow her recommendation when she believes that r_{-i} is distributed according to $\Pr^{\mu}(r_{-i}|r_i)$. ■

The next lemma says that any payoff vector that can be attained by a strategy satisfying weak incentive compatibility on path is (virtually) attainable in strict equilibrium.

Lemma 2 *With mediated perfect monitoring, suppose there exists a mediator's strategy μ such that each player has a weak incentive to follow her recommendation on-path, when she is minmaxed forever if she deviates: that is, for each player i and on-path history h_m^t,*

$$
\begin{aligned}
&(1-\delta)u_i(r_i, \mu|_{r_i}(h_m^t)) + \delta\mathbb{E}\left[\sum_{\tau=t+1}^{\infty} \delta^{\tau-1} u_i(\mu(h_m^{\tau})) \mid h_m^t, r_i\right] \\
\geq\ & \max_{a_i \in A_i}(1-\delta)u_i(a_i, \mu|_{r_i}(h_m^t)) + \delta \min_{\alpha_{-i}\in\Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}). \qquad (3)
\end{aligned}
$$

Here, $\mu|_{r_i}(h_m^t)$ is the distribution of $r_{-i,t}$ conditional on h_m^t and $r_{i,t} = r_i$. Let v be the payoff vector when each player obeys μ. Suppose also that there exists a strict full-support equilibrium. (For example, such an equilibrium exists if $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ *by Lemma 1.) Then $v \in \overline{E_{\text{med}}(\delta)}$.*

Proof. Arbitrarily fix μ and a 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} 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). On the other hand, 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 full-support equilibrium, it is always strictly optimal for each player i to follow her recommendation. ■

Note that Lemmas 1 and 2 hold for any $|I| \geq 2$.

4.2 Proof of Theorem 1: Construction

We wish to establish (1) for every Pareto weight $\lambda \in \Lambda_+$. Note that Lemma 1 implies

$$\bigcap_{i \in I} W_i \cap u(A) \subseteq \overline{E_{\text{med}}(\delta)}.$$

Therefore, for every Pareto weight $\lambda \in \Lambda_+$, if there exists $v \in \arg\max_{v' \in \mathcal{E}(\delta)} \lambda \cdot v'$ such that $v \in \bigcap_{i \in I} W_i \cap u(A)$, then there exists $v^* \in \overline{E_{\text{med}}(\delta)}$ such that $\lambda \cdot v \leq \lambda \cdot v^*$, as desired.

Hence, we are left to consider the case where

$$\arg\max_{v' \in \mathcal{E}(\delta)} \lambda \cdot v' \cap \left\{ \bigcap_{i \in I} W_i \cap u(A) \right\} = \emptyset. \tag{4}$$

Since we consider two-player games, we can order $\lambda \in \Lambda_+$ as follows: $\lambda \leq \lambda'$ 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

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

Note that the assumption that $\bigcap_{i \in I} W_i \cap u(A) \neq \emptyset$ implies that $\bar{w}^i \in \bigcap_{i \in I} W_i \cap u(A)$. 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 (non-empty) set of Pareto weight λ^i such that $\bar{w}^i \in \arg\max_{v \in u(A)} \lambda^i \cdot v$:

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

By (4) and convexity of $u(A)$, either $\lambda < \lambda^1$ for each $\lambda^1 \in \Lambda^1$ or $\lambda > \lambda^2$ for each $\lambda^2 \in \Lambda^2$. See Figure 1. We focus on the case with $\lambda > \lambda^2$. (The proof for the case with $\lambda < \lambda^1$ is symmetric and thus omitted.)

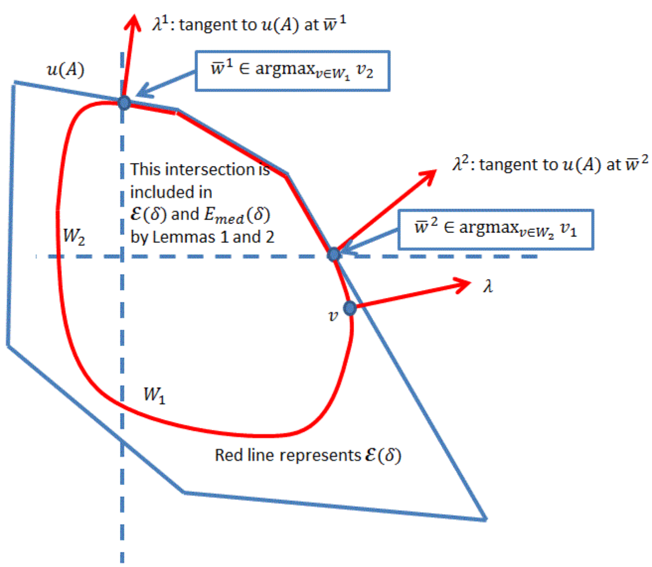

Figure 1: Setup for the construction.

Fix $v \in \arg\max_{v' \in \mathcal{E}(\delta)} \lambda \cdot v'$. Let $(\mu, (\sigma_i)_{i \in I})$ be an equilibrium that attains v. By Lemma 2, it suffices to construct a mediator's strategy μ^* yielding payoffs v^* such that (3) holds and $\lambda \cdot v \leq \lambda \cdot v^*$. The rest of the proof constructs such a strategy.

Given μ, for each on-path history of player 2's recommendations, denoted by $r_2^t = (r_{2,\tau})_{\tau=1}^{t-1}$, let $w^\mu(r_2^t)$ be the continuation payoff from period t onward conditional on r_2^t:

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

and let $\Pr^{\mu}(\cdot|r_2^t)$ be the conditional distribution of recommendations in period t. We start by defining value functions w_2^* and w_2, which map histories of player 2's recommendations to continuation payoffs for player 2. Given $w_2(r_2^{t+1})$, define $w_2^*(r_2^t)$ by

$$w_2^*(r_2^t) = (1-\delta)\, u_2\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w_2(r_2^{t+1})|r_2^t\right]. \tag{5}$$

On the other hand, given $w_2^*(r_2^t)$, define $w_2(r_2^t)$ by

$$w_2\left(r_2^t\right) = 1_{\left\{w^{\mu}(r_2^t)\in W_1\right\}} \left\{p(r_2^t) w_2^*(r_2^t) + \left(1-p(r_2^t)\right)\bar{w}_2^1\right\} + 1_{\left\{w^{\mu}(r_2^t)\notin W_1\right\}} \bar{w}_2^1, \tag{6}$$

where, when $w^{\mu}(r_2^t) \in W_1$, $p(r_2^t)$ is the largest number in $[0,1]$ such that

$$p(r_2^t) w_2^*(r_2^t) + \left(1-p(r_2^t)\right)\bar{w}_2^1 \geq w_2^{\mu}(r_2^t). \tag{7}$$

That is, if $w_2^*\left(r_2^t\right) \geq w_2^{\mu}(r_2^t)$, then $p(r_2^t) = 1$; and otherwise, since $w^{\mu}(r_2^t) \in W_1$ implies that $w_2^{\mu}(r_2^t) \leq \bar{w}_2^1$, $p(r_2^t) \in [0,1]$ solves

$$p(r_2^t) w_2^*(r_2^t) + \left(1-p(r_2^t)\right)\bar{w}_2^1 = w_2^{\mu}(r_2^t).$$

(Intuitively, the term $1_{\left\{w^{\mu}(r_2^t)\notin W_1\right\}} \bar{w}_2^1$ in (6) indicates that, instead of promising player 2 a continuation payoff outside of W_1, we promise him his best continuation payoff within W_1, namely $\bar{w}^1$. Anticipating that this replacement will occur in future may reduce $w_2^*\left(r_2^t\right)$ below player 2's original value $w_2^{\mu}(r_2^t)$. However, (7) ensures that, by promising $\bar{w}^1$ with high enough probability, player 2's value $w_2\left(r_2^t\right)$ does not fall below $w_2^{\mu}(r_2^t)$.)

We prove that there is a unique pair of value functions w_2 and w_2^* satisfying (5) and (6) for all histories $r_2^t \in A_2^{t-1}$. In particular, any such value function w_2 is a fixed point of the following operator F: Given w_2, define $w_2^*(w_2)$ according to (5). Given $w_2^*(w_2)$, define $p(w_2)$ according to (7), and then define $F(w_2)$ from $w_2^*(w_2)$ and $p(w_2)$ according to (6). We show that the operator F has a unique fixed point. It suffices to show that F is a contraction.

Lemma 3 *For all w_2 and $\tilde{w}_2$, we have $\|F(w_2) - F(\tilde{w}_2)\| \leq \delta \|w_2 - \tilde{w}_2\|$, where $\|w_2 - \tilde{w}_2\| \equiv \sup_{r_2^t} \|w_2(r_2^t) - \tilde{w}_2(r_2^t)\|$.*

Proof. By (5), $\|w_2^*(w_2) - w_2^*(\tilde{w}_2)\| \leq \delta \|w_2 - \tilde{w}_2\|$. By (6),

$$\begin{aligned} \left|F(w_2)\left(r_2^t\right) - F(\tilde{w}_2)\left(r_2^t\right)\right| &= 1_{\left\{w^\mu(r_2^t) \in W_1\right\}} \left| \begin{array}{c} \{p(w_2)(r_2^t) w_2^*(w_2)(r_2^t) + (1 - p(w_2)(r_2^t))\, \bar{w}_2^1\} \\ - \{p(\tilde{w}_2)(r_2^t) w_2^*(\tilde{w}_2)(r_2^t) + (1 - p(\tilde{w}_2)(r_2^t))\, \bar{w}_2^1\} \end{array} \right| \\ &\leq \|w_2^*(w_2) - w_2^*(\tilde{w}_2)\|. \end{aligned}$$

Combining these inequalities yields $\|F(w_2) - F(\tilde{w}_2)\| \leq \delta \|w_2 - \tilde{w}_2\|$. ■

We are now ready to construct the mediator's on-path strategy μ^*. The mediator has two states, $\omega_t \in \{S_1, S_2\}$. Recommendations are as follows:

1. In state S_1, at history $r^t = (r_\tau)_{\tau=1}^{t-1}$, the mediator recommends r_t according to $\Pr^\mu(\cdot|r_2^t)$.

2. In state S_2, the mediator recommends r_t according to some $\alpha^1 \in \Delta(A)$ such that $u(\alpha^1) = \bar{w}^1 \in W_1$.

The initial state is $\omega_1 = S_1$. 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:

1. If $w^\mu(r_2^{t+1}) \notin W_1$, then $\omega_{t+1} = S_2$ with probability one.

2. If $w^\mu(r_2^{t+1}) \in W_1$, then $\omega_{t+1} = S_2$ with probability $1 - p(r_2^{t+1})$.

Under strategy μ^*, by (5) and (6), $w_2^*(r_2^t)$ is player 2's expected continuation payoff at history r_2^t conditional on the event $\omega_t = S_1$, and $w_2(r_2^t)$ is player 2's continuation payoff at history r_2^t and state $\omega_{t-1} = S_1$. Similarly, defining $w_1^*(r_2^t)$ and $w_1(r_2^t)$ as player 1's expected continuation payoffs at history r_2^t conditional on the events $\omega_t = S_1$ and $\omega_{t-1} = S_1$, respectively, the payoff vectors $w^*(r_2^t)$ and $w(r_2^t)$ satisfy

$$w^*(r_2^t) = (1 - \delta)\, u\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w(r_2^{t+1}) | r_2^t\right] \tag{8}$$

and

$$w\left(r_2^t\right)=1_{\left\{w^\mu(r_2^t)\in W_1\right\}}\left\{p(r_2^t)w^*(r_2^t)+\left(1-p(r_2^t)\right)\bar{w}^1\right\}+1_{\left\{w^\mu(r_2^t)\notin W_1\right\}}\bar{w}^1. \tag{9}$$

In particular, recalling that $\omega_0=S_1$ and $v=w^\mu(\emptyset)\in W_1$, the ex ante payoff vector v^* is given by

$$v^*=w(\emptyset)=p(\emptyset)w^*(\emptyset)+(1-p(\emptyset))\,\bar{w}^1.$$

We prove the following key lemma in the appendix.

Lemma 4 *For all $t\geq 1$, if $w^\mu(r_2^t)\in W_1$, then $p(r_2^t)w^*(r_2^t)+(1-p(r_2^t))\,\bar{w}^1$ Pareto dominates $w^\mu(r_2^t)$.*

Here is a graphical explanation of Lemma 4: By (8), $w^*(r_2^t)-w^\mu(r_2^t)$ is parallel to $w(r_2^{t+1})-w^\mu(r_2^{t+1})$. To evaluate this difference, consider (9) for period $t+1$. The term $1_{\left\{w^\mu(r_2^{t+1})\notin W_1\right\}}\bar{w}^1$ means that we construct $w(r_2^{t+1})$ by replacing some continuation payoff not included in W_1 with $\bar{w}^1$. Hence, $w(r_2^{t+1})-w^\mu(r_2^{t+1})$ (and thus $w^*(r_2^t)-w^\mu(r_2^t)$) is parallel to $\bar{w}^1-\hat{w}(r_2^{t+1})$ for some $\hat{w}(r_2^{t+1})\in u(A)\setminus W_1$. See Figure 2 for an illustration.

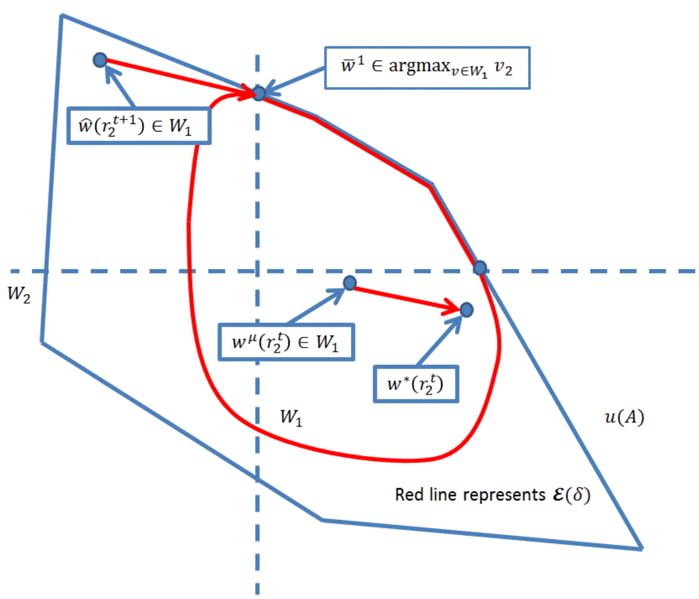

Figure 2: The vector from $w^\mu(r_2^t)$ to $w^*(r_2^t)$ is parallel to the one from $\hat{w}(r_2^{t+1})$ to $\bar{w}^1$.

Recall that $p(r_2^t)$ is determined by (7). Since the vector $w^*(r_2^t) - w^\mu(r_2^t)$ is parallel to $\bar{w}^1 - \hat{w}(r_2^{t+1})$ for some $\hat{w}(r_2^{t+1}) \in u(A) \setminus W_1$ and $u(A)$ is convex, we have $w_1^*(r_2^t) \geq w_1^\mu(r_2^t)$. Hence, if we take $p(r_2^t)$ so that the convex combination of $w_2^*(r_2^t)$ and $\bar{w}_2^1$ is equal to $w_2^\mu(r_2^t)$, then player 1 is better off compared to $w_1^\mu(r_2^t)$. See Figure 3.

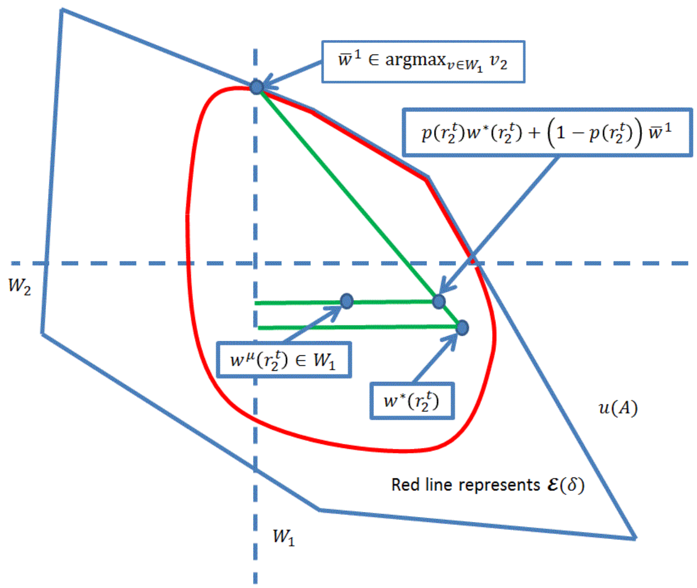

Figure 3: $p(r_2^t)w^*(r_2^t) + (1 - p(r_2^t))\,\bar{w}^1$ and $w^\mu(r_2^t)$ have the same value for player 2.

Given Lemma 4, we show that μ^* satisfies (3) for both players and $\lambda \cdot v \leq \lambda \cdot v^*$.

1. Obedience for player 1: It suffices to show that for $w(r_2^t) \in W_1$ for every on-path history r_2^t. If $\omega_t = S_2$, then $w(r_2^t) = \bar{w}^1 \in W_1$. If $\omega_t = S_1$, then $w(r_2^t)$ is given by (9). In this case, if $w^\mu(r_2^t) \in W_1$, then, by Lemma 4, $p(r_2^t)w^*(r_2^t) + (1 - p(r_2^t))\,\bar{w}^1$ Pareto dominates $w^\mu(r_2^t) \in W_1$, so $w(r_2^t) \in W_1$. If $w^\mu(r_2^t) \notin W_1$, then $w(r_2^t) = \bar{w}^1 \in W_1$. Hence, in any case $w(r_2^t) \in W_1$.

2. Obedience for player 2: If $\omega_t = S_2$ or if both $\omega_t = S_1$ and $w^\mu(r_2^t) \notin W_1$, then $w(r_2^t) = \bar{w}^1 \in W_2$, so (3) holds. If $\omega_t = S_1$ and $w^\mu(r_2^t) \in W_1$, then $p(r_2^t)w^*(r_2^t) + (1 - p(r_2^t))\,\bar{w}^1 = w_2^\mu(r_2^t)$ by (7). In this case, player 2's continuation payoff at r_2^t under μ^* is the same as it is under μ. As player 2's continuation payoff after a deviation under μ is at least

his minmax payoff, incentive compatibility at $r_2^t = (r_{2,\tau})_{\tau=1}^{t-1}$ under μ implies (3).

3. $\lambda \cdot v \leq \lambda \cdot v^*$: Immediate from Lemma 4 with $t = 1$.

5 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 $\text{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$. As our goal is giving a recursive upper bound on $E(\delta, p)$, it remains to recursively characterize $\overline{E_{\text{med}}(\delta)}$.

The idea is to replace APS's generating operator B with a "minmax-threat" generating operator $\tilde{B}$. We begin with some definitions. Given a correlated action profile $\alpha \in \Delta(A)$, let $\alpha | a_i$ be the marginal distribution on A_{-i} conditional on a_i.

Definition 1 *For any set $V \subseteq \mathbb{R}^{|I|}$, a correlated action profile $\alpha \in \Delta(A)$ is* minmax-threat enforceable *on V if there exists a mapping $\gamma : A \to V$ such that, for each player i and action $a_i \in \operatorname{supp} \alpha_i$,*

$$\begin{aligned} & E_{\alpha|a_i}\left[(1-\delta) u_i(a_i, a_{-i}) + \delta \gamma(a_i, a_{-i})\right] \\ \geq\ & \max_{a_i' \in A_i} E_{\alpha|a_i}\left[(1-\delta) u_i(a_i', a_{-i})\right] + \delta \min_{\alpha_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}). \end{aligned}$$

Definition 2 *A payoff vector $v \in \mathbb{R}^{|I|}$ is* minmax-threat decomposable *on V if there exists a correlated action profile $\alpha \in \Delta(A)$ which is minmax-threat enforced on V by some mapping γ such that*

$$v = E_\alpha\left[(1-\delta) u(a) + \delta \gamma(a)\right].$$

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

We will show that the following algorithm recursively computes $\overline{E_{\text{med}}(\delta)}$: let

$$\begin{aligned} W^1 &= u(A), \\ W^n &= \tilde{B}\left(W^{n-1}\right) \text{ for } n > 1, \\ W^\infty &= \lim_{n\to\infty} W^n. \end{aligned}$$

Note that $\lim_{n\to\infty} W^n$ exists because $\tilde{B}(u(A)) \subseteq u(A)$ and $\tilde{B}$ is a monotone operator (as if v is minmax-threat decomposable on V, then it is also minmax-threat decomposable on any V' containing V). We will also use the fact that $\tilde{B}$ preserves compactness: if $V \subseteq \mathbb{R}^{|I|}$ is compact, then so is $\tilde{B}(V)$.[9]

Our result is the following:

Theorem 2 *If* $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$, *then* $\overline{E_{\text{med}}(\delta)} = W^\infty$.

Following APS, the proof relies on the notion of self-generation.

Definition 3 *A set* $V \subseteq \mathbb{R}^{|I|}$ *is* minmax-threat self-generating *if* $V \subseteq \tilde{B}(V)$.

For the next two lemmas, assume $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$.

Lemma 5 *If a set* $V \subseteq \mathbb{R}^{|I|}$ *is bounded and minmax-threat self-generating, then* $\tilde{B}(V) \subseteq \overline{E_{\text{med}}(\delta)}$.

Proof. By the standard proof that the set of payoffs decomposable on a bounded, self-generating set is contained in the sequential equilibrium payoff set (Theorem 1 of APS, Proposition 7.3.1 of Mailath and Samuelson (2006)), for every $v \in \tilde{B}(V)$, there exists an on-path recommendation strategy μ that yields payoff v and satisfies (3). Under the assumption that $\operatorname{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$, the conclusion of the lemma follows from Lemma 2. ■

Lemma 6 $\overline{E_{\text{med}}(\delta)} = \tilde{B}\left(\overline{E_{\text{med}}(\delta)}\right)$.

[9]The proof is as in Lemma 1 of APS or Lemma 7.3.2 of Mailath and Samuelson (2006).

Proof. By Lemma 5, it suffices to show that $\overline{E_{\text{med}}(\delta)}$ is bounded and satisfies $\overline{E_{\text{med}}(\delta)} \subseteq \tilde{B}\left(\overline{E_{\text{med}}(\delta)}\right)$. Boundedness is immediate as $\overline{E_{\text{med}}(\delta)} \subseteq u(A)$, so we need only show that $\overline{E_{\text{med}}(\delta)} \subseteq \tilde{B}\left(\overline{E_{\text{med}}(\delta)}\right)$.

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, at any history h_i^t consistent with player i's opponents' strategies, player i's continuation payoff must be at least $\min_{\alpha_{-i}\in\Delta(A_{-i})}\max_{a_i} u_i(a_i,\alpha_{-i})$. 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 int $\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$, we have $E_{\text{med}}^{\text{weak}}(\delta) \subseteq \overline{E_{\text{med}}(\delta)}$.

Now, for any $v \in E_{\text{med}}(\delta)$, let $\mu(\emptyset)$ be a corresponding equilibrium period 1 recommendation 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 $\min_{\alpha_{-i}\in\Delta(A_{-i})}\max_{a_i} u_i(a_i,\alpha_{-i})$ (as the resulting history is consistent with player i's opponents' equilibrium strategies). Hence, we have

$$
\begin{aligned}
& E_{\mu|a_i}\left[(1-\delta)\,u_i(a_i,a_{-i}) + \delta w_i(a_i,a_{-i})\right] \\
\geq\ & \max_{a_i'\in A_i} E_{\mu|a_i}\left[(1-\delta)\,u_i(a_i',a_{-i})\right] + \delta \min_{\alpha_{-i}\in\Delta(A_{-i})}\max_{a_i\in A_i} u_i(a_i,\alpha_{-i}),
\end{aligned}
$$

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 i's 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 $\mu(\emptyset)$ and continuation payoff function w, so in particular $v \in \tilde{B}\left(\overline{E_{\text{med}}(\delta)}\right)$.

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

Lemma 7 *W^{∞} is minmax-threat self-generating: $W^{\infty} \subseteq \tilde{B}(W^{\infty})$.*

Proof. As $\tilde{B}$ preserves compactness, this follows by standard arguments. For example, the proof of Proposition 7.3.3 of Mailath and Samuelson (2006) applies verbatim (see also Theorem 5 of APS). ■

Proof of Theorem 2. As $\tilde{B}$ is monotone and $\overline{E_{\text{med}}(\delta)} \subseteq u(A)$ is a fixed point of $\tilde{B}$ (by Lemma 6), it follows that $\overline{E_{\text{med}}(\delta)} \subseteq W^n$ for all n. Hence, $\overline{E_{\text{med}}(\delta)} \subseteq W^{\infty}$. The reverse inclusion is immediate from Lemmas 5 and 7, as W^{∞} is bounded. ■

Combining Theorems 1 and 2 yield our main conclusion: in two-player games satisfying $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$, the equilibrium payoff set with mediated perfect monitoring is a recursive upper bound on the equilibrium payoff set with any imperfect private monitoring structure.

6 The Upper Bound in an Example

In this section, we illustrate how our operator $\tilde{B}$ works to exclude payoff vectors from the set $\overline{E_{\text{med}}(\delta)}$, and therefore from the equilibrium payoff set with 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	$0,80$	$15,90$	$35,85$	$65,65$

Note that L (low price) is myopically optimal, W (price war) is costly but hurts the other firm more than L does, and (H,H) maximizes the sum of the firms' profits. The feasible payoff set is given by

$$u(A) = \mathrm{co}\left\{(15,15),(80,0),(85,15),(80,35),(65,65),(35,80),(15,85),(0,80)\right\}.$$

Theorems 1 and 2 imply that the operator $\tilde{B}$ may be used to upper-bound the equilibrium payoff set with private monitoring only if $\mathrm{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$. One can check that a firm's minmax payoff $\underline{u}_i$ is 25 and a firm's maximum deviation gain d_i is also 25. Hence, $\mathrm{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ if and only if δ is greater than $\bar{\delta} = \frac{5}{13}$, which is the solution to

$$\frac{1-\delta}{\delta} \underbrace{25}_{\text{maximum deviation gain}} = \underbrace{65}_{\text{best symmetric payoff}} - \underbrace{25}_{\text{minmax payoff}}.$$

We therefore fix $\delta = \frac{1}{2} > \frac{5}{13}$.

We now illustrate how our algorithm works. In particular, we set $W^1 = u(A)$, and we show how to calculate $\tilde{B}(W^1)$. By Theorems 1 and 2, any Pareto-efficient payoff profile not included in $\tilde{B}(W^1)$ is not included in $\overline{E_{\mathrm{med}}(\delta)}$, or in $E(\delta, p)$ for any p.

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

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

Hence, Lemma 1 and Theorem 2 imply that

$$\begin{aligned} & \{v \in u(A) : v_i \geq 50 \text{ for each } i\} \\ = \ & \mathrm{co}\left\{(50,50),(75,50),(50,75),(65,65)\right\} \subseteq W^n \end{aligned}$$

for each n. As in Section 4.2, let λ^1 be the tangential vector of $u(A)$ at $(50,75)$, and let λ^2 be the one at $(75,50)$. Hence,

$$\lambda^1 = \begin{pmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{pmatrix}, \lambda^2 = \begin{pmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix}.$$

As $\tilde{B}(W^1)$ is convex, to calculate its Pareto frontier it suffices to find $v \in \arg\max_{v' \in \tilde{B}(W^1)} \lambda \cdot$

v' for each $\lambda \in \Lambda_+$. For λ with $\frac{\lambda_1^1}{\lambda_2^1} \leq \frac{\lambda_1}{\lambda_2} \leq \frac{\lambda_1^2}{\lambda_2^2}$, v is included in

$$\text{co}\left\{(50,50),(75,50),(50,75),(65,65)\right\},$$

as in Lemma 1. Hence, we focus on Pareto weights λ with $\frac{\lambda_1}{\lambda_2} > \frac{\lambda_1^2}{\lambda_2^2}$, as in the proof of Theorem 1. (The calculation for λ with $\frac{\lambda_1}{\lambda_2} < \frac{\lambda_1^1}{\lambda_2^1}$ is symmetric.)

For example, consider $\lambda = (1,0)^\top$: that is, we maximize firm 1's payoff. Since we put a higher weight on firm 1's payoffs, it is natural to conjecture that firm 1's incentive constraint is not binding. Hence, we 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. On the other hand, given firm 2's strategy, firm 1's own payoff increases as firm 1 increases its price from W to L and decreases as it increases the price from L to M or H. Hence, in order 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

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

where w_2 is firm 2's continuation payoff. That is, $w_2 \geq 50$.

By feasibility, $w_2 \geq 50$ implies that $w_1 \leq 75$. Hence, conditional on $r_2 = H$, the best payoff for firm 1 that is minmax-threat decomposable is

$$(1-\delta)\begin{pmatrix} 90 \\ 15 \end{pmatrix} + \delta \begin{pmatrix} 75 \\ 50 \end{pmatrix} = \begin{pmatrix} \frac{165}{2} \\ \frac{65}{2} \end{pmatrix}.$$

Since $\frac{165}{2}$ 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 $75 \geq \underline{w}_1$, firm 1's incentive constraint is not binding.

Therefore, the set $\overline{E_{\text{med}}(\delta)}$ does contain any payoff vectors v with $v_1 > \frac{165}{2}$. On the other hand, the set of feasible and individually rational payoffs is

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

so we have eliminated some feasible and individually rational payoffs. Thus, our algorithm gives a substantive upper bound on firm 1's payoff.

7 Extensions

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

7.1 What if $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) = \emptyset$?

The assumption that $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ guarantees that all action profiles are supportable in equilibrium (Lemma 1), which in turn implies that deviators can be held to their correlated minmax payoffs (Lemma 2). This fact plays a key role both in showing that $\overline{E_{\text{med}}(\delta)}$ is an upper bound on $E(\delta, p)$ for any private monitoring structure p and in recursively characterizing $\overline{E_{\text{med}}(\delta)}$. However, the assumption that all action profiles are supportable is restrictive: for instance, this assumption will often fail in models of Bertrand oligopoly (where very high prices are not enforceable) or repeated agency (where very high effort is not enforceable). The assumption that $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ also 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 non-trivial payoff bounds for all Pareto weights.

To address these concerns, this subsection shows that even if $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) = \emptyset$, $\overline{E_{\text{med}}(\delta)}$ may still be an upper bound on $E(\delta, p)$ for any private monitoring structure p, and $\overline{E_{\text{med}}(\delta)}$ 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 int $\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ holds with respect to the subset of action profiles that are supportable.

Recall that $\mathcal{E}(\delta)$ is the equilibrium payoff set with the monitoring structure where a mediator observes the recommendation profile r_t and action profile a_t in each period t, while each player i only observes her own recommendation $r_{i,t}$ and her own action $a_{i,t}$. Denote this monitoring structure by p^*. We want to provide a sufficient condition for the Pareto frontier of $\overline{E_{\text{med}}(\delta)}$ to dominate the Pareto frontier of $\mathcal{E}(\delta)$, and to characterize $E_{\text{med}}(\delta)$.

Let supp(δ) be the set of supportable actions with monitoring structure p^*:

$$\text{supp}(\delta) = \left\{ a \in A : \begin{array}{c} \text{with monitoring structure } p^*, \\ \text{there exist an equilibrium strategy } \mu \\ \text{and history } h_m^t \text{ with } a \in \text{supp}(\mu(h_m^t)) \end{array} \right\}.$$

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

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

$$\begin{aligned} & (1-\delta)\, u_i(a_i, \alpha_{-i}) + \delta \max_{a_i \in A_i, \bar{a}_{-i} \in \bar{A}_{-i}} u_i\left(a_i, \bar{a}_{-i}\right) \\ \geq\ & (1-\delta) \max_{\hat{a}_i \in A_i} u_i(\hat{a}_i, \alpha_{-i}) + \delta \min_{\hat{\alpha}_{-i} \in \Delta\left(\bar{A}_{-i}\right)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}). \end{aligned} \tag{10}$$

Let $S\left(\bar{A}\right) = \prod_{i\in I} S_i\left(A_i\right) \subseteq \bar{A}$. Let $\mathcal{A}^1 = A$, let $\mathcal{A}^n = S\left(\mathcal{A}^{n-1}\right)$ 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 $S\left(\bar{A}\right)$ is defined by a finite number of linear inequalities.

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

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

be the lowest continuation payoff such that player i does not want to deviate 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

$$\bar{W}_i = \left\{ w \in \mathbb{R}^N : w_i \geq \min_{\alpha_{-i} \in \Delta\left(\mathcal{A}^\infty_{-i}\right)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) + \frac{1-\delta}{\delta} \max_{r \in \mathcal{A}^\infty, a_i \in A_i} \{u_i(a_i, r_{-i}) - u_i(r)\} \right\}.$$

We show the following.

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

$$\operatorname{int}\left(\bigcap\nolimits_{i \in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset \tag{11}$$

in the topology induced from $u(\mathcal{A}^\infty)$, then for every private monitoring structure p and every non-negative 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. Given that $\operatorname{supp}(\delta) = \mathcal{A}^\infty$, the proof is analogous to the proof of Theorem 1, everywhere replacing $u(A)$ with $u(\mathcal{A}^\infty)$ and replacing "full support" with "full support within $\mathcal{A}^\infty$." The proof that $\operatorname{supp}(\delta) = \mathcal{A}^\infty$ whenever $\operatorname{int}\left(\bigcap_{i \in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset$ is deferred to the appendix. ■

Note that Proposition 2 only improves on Theorem 1 at "low" discount factors: for a high enough discount factor, $A = S(A) = \mathcal{A}^\infty$, so (11) reduces to $\operatorname{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$. However, as we have seen, $\operatorname{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right)$ can be empty only for low discount factors, so the low discount factor case is precisely the one where an improvement is needed.[10]

In order to be able to use Proposition 2 to give a recursive upper bound on $E(\delta, p)$ when $\operatorname{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$, we must characterize $\overline{E_{\text{med}}(\delta)}$ under (11). Our earlier

[10]To be clear, it is possible for $\operatorname{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right)$ to be empty while $\mathcal{A}^\infty = A$. Theorem 1 and Proposition 2 only give sufficient conditions: we are not claiming that they cover every possible case.

characterization generalizes easily. In particular, the following definitions are analogous to Definitions 1 and 2.

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

$$\begin{aligned} & E_{\alpha|a_i}\left[(1-\delta)\,u_i\left(a_i, a_{-i}\right) + \delta\gamma\left(a_i, a_{-i}\right)\right] \\ \geq\ & \max_{a_i' \in A_i} E_{\alpha|a_i}\left[(1-\delta)\,u_i\left(a_i', a_{-i}\right)\right] + \delta \min_{\alpha_{-i} \in \Delta(\text{supp}(\delta))} \max_{a_i \in A_i} u_i\left(a_i, \alpha_{-i}\right). \end{aligned}$$

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

$$v = E_{\alpha}\left[(1-\delta)\,u(a) + \delta\gamma(a)\right].$$

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

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

Proposition 3 *If $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^{\infty}\right) \neq \emptyset$, then $\overline{E_{\text{med}}(\delta)} = W^{\mathcal{A}^{\infty},\infty}$.*

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

As an example of how Propositions 2 and 3 can be applied, one can check that applying the operator S in the Bertrand example in Section 6 for any $\delta \in \left(\frac{1}{4}, \frac{5}{18}\right)$ yields $\mathcal{A}^{\infty} = \{W, L, M\} \times \{W, L, M\}$—ruling out the efficient action profile (H, H)—and $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^{\infty}\right) \neq \emptyset$. We can then compute $\overline{E_{\text{med}}(\delta)}$ by applying the operator $\tilde{B}^{\mathcal{A}^{\infty}}$, just as we as we applied $\tilde{B}$ in Section 6.

Finally, we mention that $\overline{E_{\text{med}}(\delta)}$ can be characterized recursively even if $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^{\infty}\right) = \emptyset$. This fact is not directly relevant for the current paper, so we omit the proof. It is available from the authors upon request.

7.2 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 with private monitoring.

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 $p_t(\cdot|a^t, z^t)$. (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.[11] In the space of such *stochastic* private monitoring structures, 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 stochastic private monitoring structure p, because the equilibrium payoff set with the universal monitoring structure p^*, $\mathcal{E}(\delta)$, remains an upper bound on $E(\delta, p)$; and the bound is tight because mediated perfect monitoring is itself a particular stochastic private monitoring structure.

Second, maintaining the assumption that the monitoring structure is "repeated" (i.e., non-stochastic), 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.[12] For this result, we also need to slightly generalize our definition of a private monitoring structure by letting the players get signals before the first round of play. This seems innocuous, especially if we take the perspective of an an outside observer who does not know the game's start date. Equivalently, the players have access to a mediator at the beginning of the game only. The monitoring structure is required to be repeated thereafter. We call such a monitoring structure a *private monitoring structure with ex ante correlation.*

[11] 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."

[12] This idea is as in the literature on implementing correlated equilibria without a mediator (see Forges (2009) for a survey).

Proposition 4 *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 with finitely many rounds of public cheap talk before each round of play. If $|I| = 2$, there exists a private monitoring structure with ex ante correlation p such that $E_{\text{talk}}(\delta, p) = E_{\text{med}}(\delta)$.*

Proof. See appendix. ■

If instead one insists on repeated monitoring and does not allow communication, then we believe that there are some games in which our bound is not tight, in that there are points in $\overline{E_{\text{med}}(\delta)}$ which are not attainable in equilibrium for any repeated private monitoring structure. We leave this as a conjecture.[13]

7.3 What if There are More Than Two Players?

The condition that $\text{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$ 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 $\text{int}\left(\bigcap_{i \in I} W_i \cap u(A)\right) \neq \emptyset$ but $\max_{v \in E(\delta,p)} \lambda \cdot v > \max_{v \in \overline{E_{\text{med}}(\delta)}} \lambda \cdot v$ for some private monitoring structure p and some non-negative Pareto weight $\lambda \in \Lambda_+$.*

Proof. We give an example in the appendix. ■

To see where the proof of Theorem 1 breaks down with $|I| > 2$, recall that the proof is based on 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

[13] 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 $\overline{E_{\text{med}}(\delta)}$ which 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 non-trivial conjecture is that the bound is still not tight when ex ante correlation is allowed (but communication is not).

she necessarily knows in any private monitoring structure). On the other hand, 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 2 that $\text{supp}(\delta) = \mathcal{A}^\infty$ whenever $\text{int}\left(\bigcap_{i \in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset$ does not rely on $|I| = 2$. Thus, when $\text{int}\left(\bigcap_{i \in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset$, we can characterize the set of supportable actions for any number of players. This is sometimes already enough to imply a non-trivial upper bound on payoffs.

Second, Proposition 6 below shows that if a payoff vector $v \in \text{int}\,(u\,(A))$ satisfies $v_i > \underline{u}_i + \frac{1-\delta}{\delta} d_i$ for all $i \in I$, then $v \in \overline{E_{\text{med}}(\delta)}$. This shows that private monitoring cannot do "much" better than mediated perfect monitoring when the players are at least moderately patient.

Finally, suppose there is a player i whose opponents $-i$ all have identical payoff functions. 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.

7.4 The Folk Theorem with Mediated Perfect Monitoring

The point of this paper is bounding the equilibrium payoff set with private monitoring at a fixed discount factor. Nonetheless, it is worth pointing out that—as a corollary of our results—a strong version of the folk theorem holds with mediated perfect monitoring, without any conditions on the feasible payoff set (e.g., full-dimensionality, non-equivalent utilities), and with a rate of convergence of $1 - \delta$. One reason why this result may be of interest is that it shows that, if players are fairly patient, they cannot do "much" better with private monitoring than with mediated perfect monitoring, even if the sufficient conditions for Theorem 1 fail. Another reason is that the mediator can usually be replaced by unmediated communication among the players (we spell out exactly when below), and it seems worth knowing that the folk theorem holds with unmediated communication without conditions on the feasible payoff set.

Recall that $\underline{u}_i$ is player i's correlated minmax payoff and that d_i is player i's greatest possible deviation gain. Let $\underline{u} = (u_i)_{i \in I}$ and $d = (d_i)_{i \in I}$.

Proposition 6 *If a payoff vector $v \in \text{int}\,(u\,(A))$ satisfies*

$$v > \underline{u} + \frac{1-\delta}{\delta} d, \tag{12}$$

then $v \in \overline{E_{\text{med}}(\delta)}$.

Proof. Fix $\alpha \in \Delta\,(A)$ such that $v = u\,(\alpha)$. If v satisfies (12), then $v \in \text{int}\left(\bigcap_{i \in I} W_i \cap u\,(A)\right)$ (so in particular $\text{int}\left(\bigcap_{i \in I} \cap u\,(A)\right) \neq \emptyset$), and the infinite repetition of α satisfies (3). Lemma 1 then gives $v \in \overline{E_{\text{med}}(\delta)}$. ■

The folk theorem is a corollary of Proposition 6. Recall that a payoff vector v is *strictly individually rational* (relative to correlated minmax payoffs) if $v > \underline{u}$.[14]

Corollary 1 (Folk Theorem) *For every strictly individually rational payoff vector $v \in u\,(A)$, there exists $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$ then $v \in E_{\text{med}}(\delta)$.*

Proof. It is immediate that $v \in \overline{E_{\text{med}}(\delta)}$ for high enough δ: let $\bar{\delta} = \max_{i \in I} \frac{d_i}{v_i - \underline{u}_i + d_i}$ and apply Theorem 6. The proof that taking the closure is unnecessary is in the appendix. ■

In earlier work (Sugaya and Wolitzky, 2014b), we have shown that the mediator can usually be dispensed with in Proposition 6 and Corollary 1 if players can communicate through cheap talk. In particular, this is possible unless there are exactly three players, of whom exactly two have equivalent utilities in the sense of Abreu, Dutta, and Smith (1994).

[14]The restriction to *strictly* individually rational payoff vectors cannot be dropped. In particular, the counterexample to the folk theorem due to Forges, Mertens, and Neyman (1986)—in which no payoff vector is strictly individually rational—remains valid in the presence of a mediator.

8 Conclusion

This paper gives a simple sufficient condition under which the equilibrium payoff set in a two-player repeated game with perfect monitoring and a mediator is a tight, recursive upper bound on the equilibrium payoff set in the same game with any imperfect private monitoring structure. Equivalently, under our sufficient condition, this set characterizes those payoffs that could arise in a repeated game from the perspective of an observer who does not know the monitoring structure. The sufficient condition reduces to a lower bound on the discount factor whenever there exists a feasible and strictly individual rational payoff vector. Thus, in almost all games, simple recursive methods can be used to upper-bound the equilibrium payoff set in repeated games with imperfect private monitoring at a fixed discount factor.

9 Appendix

9.1 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, 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 1. This is possible only if the probability that (T, M) is played in period $t+1$ when U is played in period t exceeds the number p such that

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

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

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

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

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

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

9.2 Proof of Proposition 1: 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 obtains with probability 1; whenever player 1 plays D, players 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 $n = 0, 1, ...$, 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.

Player 2's strategy: In each odd period $t = 2n + 1$ with $n = 0, 1, ...$, 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.

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

In an odd period, player 1's payoff from U is the solution to $v = 2 + \frac{1}{6}(6) + \frac{1}{6^2}v$. On the other hand, her payoff from D is $3 + \frac{1}{6}(0) + \frac{1}{6^2}v$. Hence, player 1 is indifferent between U and D (and clearly prefers either of these 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.

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

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

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

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.[15] Second, players' payoffs are measurable with respect to their own actions and signals. The required realized payoffs for player 2 are as follows:[16]

(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.[17]

9.3 Proof of Lemma 4

It is useful to introduce a family of auxiliary value functions $\left(w_2^T\right)_{T=1}^{\infty}$ and $\left(w_2^{*,T}\right)_{T=1}^{\infty}$, which will converge to w_2 and w_2^* pointwise in r_2^t as $T \to \infty$. For periods $t \geq T$, define

$$w_2^T(r_2^t) = w_2^{\mu}(r_2^t) \text{ and } w_2^{*,T}(r_2^{t-1}) = w_2^{\mu}(r_2^{t-1}). \tag{13}$$

On the other hand, for periods $t \leq T-1$, define $w_2^{*,T}(r_2^t)$, $p^T(r_2^t)$, and $w_2^T(r_2^t)$ given $w_2^T(r_2^{t+1})$ recursively, as follows. First, define

$$w_2^{*,T}(r_2^t) = (1-\delta)\, u_2\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w_2^T(r_2^{t+1})|r_2^t\right]. \tag{14}$$

Note that, for $t = T-1$, this definition is compatible with (13). Second, given $w_2^{*,T}(r_2^t)$, define

$$w_2^T(r_2^t) = 1_{\left\{w^{\mu}(r_2^t)\in W_1\right\}}\left\{p^T(r_2^t)w_2^{*,T}(r_2^t) + \left(1-p^T(r_2^t)\right)\bar{w}_2^1\right\} + 1_{\left\{w^{\mu}(r_2^t)\notin W_1\right\}}\bar{w}_2^1, \tag{15}$$

[15] The only non-trivial 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.

[16] Recall that player 1 can observe player 2's action perfectly.

[17] 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.

where, when $w_2^\mu(r_2^t) \in W_1$, $p^T(r_2^t)$ is the largest number in $[0,1]$ such that

$$p^T(r_2^t) w_2^{*,T}(r_2^t) + \left(1 - p^T(r_2^t)\right) \bar{w}_2^1 \geq w_2^\mu(r_2^t). \tag{16}$$

We show that $w_2^{*,T}$ converges to w_2^*.

Lemma 8 $\lim_{T\to\infty} w_2^{*,T}(r_2^t) = w_2^*(r_2^t)$ *for all* $r_2^t \in A_2^{t-1}$.

Proof. By Lemma 3, it suffices to show that $F(w_2^T) = w_2^{T+1}$. For $t \geq T+1$, (13) implies that $w_2^{*,T+1}(r_2^{t-1}) = w_2^\mu(r_2^{t-1})$. On the other hand, given w_2^T, $w_2^*(w_2^T)$ is the value calculated according to (5). Since $w_2^T(r_2^t) = w_2^\mu(r_2^t)$ by (13), we have $w_2^*(w_2^T)(r_2^{t-1}) = w_2^\mu(r_2^{t-1})$ by (5). Hence,

$$w_2^*(w_2^T)(r_2^{t-1}) = w_2^{*,T+1}(r_2^{t-1}). \tag{17}$$

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

$$w_2^{*,T+1}(r_2^t) = (1-\delta)\, u_2\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w_2^{T+1}(r_2^{t+1}) | r_2^t\right].$$

By (15),

$$\begin{aligned} w_2^{T+1}(r_2^t) &= 1_{\left\{w^\mu(r_2^t)\in W_1\right\}} \left\{p^{T+1}(r_2^t) w_2^{*,T+1}(r_2^t) + \left(1 - p^{T+1}(r_2^t)\right) \bar{w}_2^1\right\} + 1_{\left\{w^\mu(r_2^t)\notin W_1\right\}} \bar{w}_2^1 \\ &= 1_{\left\{w^\mu(r_2^t)\in W_1\right\}} \left\{p^{T+1}(r_2^t) w_2^*(w_2^T)(r_2^t) + \left(1 - p^{T+1}(r_2^t)\right) \bar{w}_2^1\right\} + 1_{\left\{w^\mu(r_2^t)\notin W_1\right\}} \bar{w}_2^1, \end{aligned}$$

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

Recall that p^{T+1} is defined by (16). On the other hand, $p(w_2^T)(r_2^t)$ is defined in (7) using $w_2^* = w_2^*(w_2^T)$. Since $w_2^{*,T+1}(r_2^t) = w_2^*(w_2^T)(r_2^t)$, we have $p^{T+1}(r_2^t) = p(w_2^T)(r_2^t)$. Hence,

$$\begin{aligned} w_2^{T+1}(r_2^t) &= 1_{\left\{w^\mu(r_2^t)\in W_1\right\}} \left\{p(w_2^T)(r_2^t) w_2^*(w_2^T)(r_2^t) + \left(1 - p(w_2^T)(r_2^t)\right) \bar{w}_2^1\right\} + 1_{\left\{w^\mu(r_2^t)\notin W_1\right\}} \bar{w}_2^1. \\ &= F(w_2^T)(r_2^t), \end{aligned}$$

as desired. ■

We also define $w_1^T(r_2^t)$ and $w_1^{*,T}(r_2^t)$ analogously to $w_1(r_2^t)$ and $w_1^*(r_2^t)$: For $t \geq T$, define

$$w_1^T(r_2^t) = w_1^\mu(r_2^t) \text{ and } w_1^{*,T}(r_2^{t-1}) = w_1^\mu(r_2^{t-1}).$$

For $t \leq T-1$, given $w_2^T(r_2^{t+1})$ and $p^T(r_2^t)$, define

$$w_1^{*,T}(r_2^t) = (1-\delta)\, u_1\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w_1^T(r_2^{t+1})|r_2^t\right]$$

and

$$w_1^T(r_2^t) = 1_{\left\{w^\mu(r_2^t)\in W_1\right\}}\left\{p^T(r_2^t)w_1^{*,T}(r_2^t) + \left(1-p^T(r_2^t)\right)\bar{w}_1^1\right\} + 1_{\left\{w^\mu(r_2^t)\notin W_1\right\}}\bar{w}_1^1.$$

As $w_2^{*,T}(r_2^t)$, $w_2^T(r_2^t)$, and $p^T(r_2^t)$ converge to $w_2^*(r_2^t)$, $w_2(r_2^t)$, and $p(r_2^t)$ by Lemma 8, it follows that $w_1^{*,T}(r_2^t)$ and $w_1^T(r_2^t)$ converge to $w_1^*(r_2^t)$ and $w_1(r_2^t)$. Hence, the following lemma implies Lemma 4:

Lemma 9 *For all $t = 1, ..., T-1$, if $w^\mu(r_2^t) \in W_1$, then $p^T(r_2^t)w^{*,T}(r_2^t) + \left(1-p^T(r_2^t)\right)\bar{w}^1$ Pareto dominates $w^\mu(r_2^t)$.*

Proof. For $t = T-1$, the claim is immediate since $w^{*,T}(r_2^t) = w^\mu(r_2^t)$ and so $p^T(r_2^t) = 1$.

Suppose that the claim holds for each period $\tau \geq t+1$. We show that it also hold for period t. By construction, $p^T(r_2^t)w_2^{*,T}(r_2^t) + \left(1-p^T(r_2^t)\right)\bar{w}_2^1 \geq w_2^\mu\left(r_2^t\right)$. Thus, it suffices to show that $p^T(r_2^t)w_1^{*,T}(r_2^t) + \left(1-p^T(r_2^t)\right)\bar{w}_1^1 \geq w_1^\mu(r_2^t)$.

Note that

$$\begin{aligned}
& w^{*,T}(r_2^t) \\
= \;& (1-\delta)\, u\left(\mu(r_2^t)\right) + \delta \mathbb{E}\left[w^T(r_2^{t+1})|r_2^t\right] \\
= \;& (1-\delta)\, u\left(\mu(r_2^t)\right) + \delta \left\{ \begin{array}{c} \sum_{r_2^{t+1}:w^\mu(r_2^{t+1})\in W_1} \Pr^\mu\left(r_2^{t+1}|r_2^t\right)\left\{p^T(r_2^{t+1})w^{*,T}(r_2^{t+1}) + \left(1-p^T(r_2^{t+1})\right)\bar{w}^1\right\} \\ + \sum_{r_2^{t+1}:w^\mu(r_2^{t+1})\notin W_1} \Pr^\mu\left(r_2^{t+1}|r_2^t\right)\bar{w}^1 \end{array} \right\},
\end{aligned}$$

while

$$w^{\mu}(r_2^t) = (1-\delta)\, u\left(\mu(r_2^t)\right) + \delta \sum_{r_2^{t+1}} \Pr^{\mu}\left(r_2^{t+1}|r_2^t\right) w^{\mu}(r_2^t).$$

Hence,

$$\begin{aligned} & w^{*,T}(r_2^t) - w^{\mu}(r_2^t) \\ = \; & \delta \left\{ \begin{array}{c} \sum_{r_2^{t+1}: w^{\mu}(r_2^{t+1}) \in W_1} \Pr^{\mu}\left(r_2^{t+1}|r_2^t\right) \left\{ p^T(r_2^{t+1}) w^{*,T}(r_2^{t+1}) + \left(1 - p^T(r_2^{t+1})\right) \bar{w}^1 - w^{\mu}(r_2^{t+1}) \right\} \\ + \sum_{r_2^{t+1}: w^{\mu}(r_2^{t+1}) \notin W_1} \Pr^{\mu}\left(r_2^{t+1}|r_2^t\right) \left\{ \bar{w}^1 - w^{\mu}(r_2^{t+1}) \right\} \end{array} \right\}. \end{aligned}$$

When $w^{\mu}(r_2^{t+1}) \in W_1$, the inductive hypothesis implies that

$$p^T(r_2^{t+1}) w^{*,T}(r_2^{t+1}) + \left(1 - p^T(r_2^{t+1})\right) \bar{w}^1 - w^{\mu}(r_2^{t+1}) \geq 0.$$

On the other hand, note that

$$\sum_{r_2^{t+1}: w^{\mu}(r_2^{t+1}) \notin W_1} \Pr^{\mu}\left(r_2^{t+1}|r_2^t\right) \left\{ \bar{w}^1 - w^{\mu}(r_2^{t+1}) \right\} = l(r_2^t)(\bar{w}^1 - \tilde{w}(r_2^t))$$

for some number $l(r_2^t) \geq 0$ and vector $\tilde{w}(r_2^t) \notin W_1$. In total, we have

$$w^{*,T}(r_2^t) = w^{\mu}(r_2^t) + l(r_2^t)(\bar{w}^1 - \hat{w}(r_2^t)) \tag{18}$$

for some number $l(r_2^t) \geq 0$ and vector $\hat{w}(r_2^t) \leq \tilde{w}(r_2^t) \notin W_1$. Since $\bar{w}_1^1 \geq \hat{w}_1(r_2^t)$, if $\bar{w}_1^1 \geq w_1^{\mu}(r_2^t)$ then (18) implies that $\min\left\{ w_1^{*,T}(r_2^t), \bar{w}_1^1 \right\} \geq w_1^{\mu}(r_2^t)$, and therefore $p^T(r_2^t) w_1^{*,T}(r_2^t) + \left(1 - p^T(r_2^t)\right) \bar{w}_1^1 \geq w_1^{\mu}(r_2^t)$. In addition, if $w^{\mu}(r_2^{t+1}) \in W_1$ with probability one, then the inductive hypothesis implies that $w_2^{*,T}(r_2^t) \geq w_2^{\mu}(r_2^t)$, and therefore $p^T(r_2^t) = 1$ and

$$\begin{aligned} p^T(r_2^t) w_1^{*,T}(r_2^t) + \left(1 - p^T(r_2^t)\right) \bar{w}_1^1 &= w_1^{*,T}(r_2^t) \\ &= w_1^{\mu}(r_2^t) + l(r_2^t)(\bar{w}_1^1 - \hat{w}_1(r_2^t)) \\ &\geq w_1^{\mu}(r_2^t). \end{aligned}$$

Hence, it remains only to consider the case where $\bar{w}_1^1 < w_1^{\mu}(r_2^t)$ and $l(r_2^t) > 0$.

In this case, take a normal vector λ^1 of the supporting hyperplane of $u(A)$ at $\bar{w}^1$. We have $\lambda_1^1 \geq 0$ and $\lambda_2^1 > 0$, and in addition (as $\hat{w}(r_2^t) \leq \tilde{w}(r_2^t) \in u(A)$ and $w^{\mu}(r_2^t) \in u(A)$)

$$\begin{aligned}\lambda^1 \cdot \left(\bar{w}^1 - \hat{w}(r_2^t)\right) &\geq 0, \\ \lambda^1 \cdot \left(\bar{w}^1 - w^{\mu}(r_2^t)\right) &\geq 0.\end{aligned}$$

As $\bar{w}_1^1 - \hat{w}_1(r_2^t) > 0$ and $\bar{w}_1^1 - w_1^{\mu}(r_2^t) < 0$, we have

$$\frac{\hat{w}_2(r_2^t) - \bar{w}_2^1}{\bar{w}_1^1 - \hat{w}_1(r_2^t)} \leq \frac{\lambda_1^1}{\lambda_2^1} \leq \frac{\bar{w}_2^1 - w_2^{\mu}(r_2^t)}{w_1^{\mu}(r_2^t) - \bar{w}_1^1}.$$

Next, by (18), the slope of the line from $w^{\mu}(r_2^t)$ to $w^{*,T}(r_2^t)$ equals the slope of the line from $\hat{w}(r_2^t)$ to $\bar{w}^1$. Hence,

$$\frac{w_2^{\mu}(r_2^t) - w_2^{*,T}(r_2^t)}{w_1^{*,T}(r_2^t) - w_1^{\mu}(r_2^t)} \leq \frac{\bar{w}_2^1 - w_2^{\mu}(r_2^t)}{w_1^{\mu}(r_2^t) - \bar{w}_1^1}.$$

In this inequality, the denominator of the left-hand side and the numerator of the right-hand side are both positive: $w_1^{*,T}(r_2^t) > w_1^{\mu}(r_2^t)$ by (18) and $l(r_2^t) > 0$, while $\bar{w}_2^1 > w_2^{\mu}(r_2^t)$ because $w^{\mu}(r_2^t) \in W_1$ and $w^{\mu}(r_2^t) \neq \bar{w}_2^1$. Therefore, the inequality is equivalent to

$$\frac{w_2^{\mu}(r_2^t) - w_2^{*,T}(r_2^t)}{\bar{w}_2^1 - w_2^{\mu}(r_2^t)} \leq \frac{w_1^{*,T}(r_2^t) - w_1^{\mu}(r_2^t)}{w_1^{\mu}(r_2^t) - \bar{w}_1^1}.$$

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

$$q w_1^{*,T}(r_2^t) + (1 - q)\,\bar{w}_1^1 = w_1^{\mu}(r_2^t).$$

Note that

$$1 - p^T(r_2^t) = \frac{w_2^{\mu}(r_2^t) - w_2^{*,T}(r_2^t)}{\bar{w}_2^1 - w_2^{\mu}(r_2^t)},$$

while

$$1-q=\frac{w_1^{*,T}(r_2^t)-w_1^{\mu}(r_2^t)}{w_1^{\mu}(r_2^t)-\bar{w}_1^1}.$$

Hence, $p^T(r_2^t)>q$. Finally, we have seen that $\bar{w}_1^1\le w_1^{\mu}(r_2^t)\le w_1^{*,T}(r_2^t)$, so we have

$$p^T(r_2^t)w_1^{*,T}\left(r_2^t\right)+\left(1-p^T(r_2^t)\right)\bar{w}_1^1\ge qw_1^{*,T}\left(r_2^t\right)+(1-q)\,\bar{w}_1^1=w_1^{\mu}(r_2^t).$$

■

9.4 Proof of Proposition 2

We wish to show that $\mathrm{supp}(\delta)=\mathcal{A}^{\infty}$ whenever $\mathrm{int}\big(\bigcap_{i\in I}\bar{W}_i\cap u\left(\mathcal{A}^{\infty}\right)\big)\neq\emptyset$. In order to characterize $\mathrm{supp}(\delta)$, it will be useful to consider ε-sequential equilibria in the repeated game with monitoring structure p^*. We say that an assessment is an ε-sequential equilibrium if it is consistent, and, for each player i and (on- or off-path) history h_i^t, player i's deviation gain at history h_i^t is no more than ε. Let

$$\mathrm{supp}_{\varepsilon}(\delta)=\left\{a\in A:\left(\begin{array}{c}\text{there exist an }\varepsilon\text{-sequential equilibrium }\mu\\ \text{such that }a\in\mathrm{supp}\left(\mu(\emptyset)\right)\end{array}\right)\right\}$$

be the on-path support of actions in the initial period in ε-sequential equilibrium with monitoring structure p^*; and let

$$\mathrm{supp}_{\varepsilon}^{t}(\delta)=\left\{a\in A:\left(\begin{array}{c}\text{there exist an }\varepsilon\text{-sequential equilibrium }\mu\text{ and history }h_m^t\\ \text{such that }a\in\mathrm{supp}\left(\mu(h_m^t)\right)\end{array}\right)\right\}$$

be the (possibly off-path) support of actions in period t.

The proof that $\mathrm{supp}(\delta)=\mathcal{A}^{\infty}$ relies on three preliminary observations. First, $\bigcap_{\varepsilon>0}\mathrm{supp}_{\varepsilon}(\delta)$ has a product structure. Second, $\mathrm{supp}_{\varepsilon}^{t}(\delta)$ is time invariant and the on-path support and off-path support coincide: $\bigcap_{\varepsilon>0}\mathrm{supp}_{\varepsilon}(\delta)=\bigcap_{\varepsilon>0}\bigcup_t\mathrm{supp}_{\varepsilon}^{t}(\delta)$. Third, the ε-sequential equilibrium correspondence is upper-hemicontinuous in ε. In particular, $\mathrm{supp}(\delta)=\bigcap_{\varepsilon>0}\bigcup_t\mathrm{supp}_{\varepsilon}^{t}(\delta)=$

$\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$.

We establish these preliminary results in turn, and then prove that $\text{supp}(\delta) = \mathcal{A}^\infty$ whenever $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u\left(\mathcal{A}^\infty\right)\right) \neq \emptyset$. We first show that $\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$ has a product structure. Let

$$\text{supp}_{i,\varepsilon}(\delta) = \left\{ a_i \in A_i : \begin{pmatrix} \text{there exists an } \varepsilon\text{-sequential equilibrium } \mu \\ \text{such that } a_i \in \text{supp}\left(\mu_i(\emptyset)\right) \end{pmatrix} \right\}$$

be the support of player i's actions, where μ_i is the marginal distribution of player i's recommendation. We show the following:

Lemma 10 *$\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$ has a product structure: $\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta) = \bigcap_{\varepsilon>0} \prod_{i\in I} \text{supp}_{i,\varepsilon}(\delta)$.*

Proof. Since $\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta) \subseteq \bigcap_{\varepsilon>0} \prod_{i\in I} \text{supp}_{i,\varepsilon}(\delta)$, we are left to show $\bigcap_{\varepsilon>0} \prod_{i\in I} \text{supp}_{i,\varepsilon}(\delta) \subseteq \bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$. To this end, fix an arbitrary ε-sequential equilibrium strategy μ. It suffices to show that, for each $\varepsilon' > \varepsilon$, there exists an ε'-sequential equilibrium strategy μ_{product} such that $\text{supp}\left(\mu_{\text{product}}(\emptyset)\right) = \prod_{i\in I} \text{supp}(\mu_i(\emptyset))$. (This is sufficient because it implies that $\bigcap_{\varepsilon>0} \prod_{i\in I} \text{supp}_{i,\varepsilon}(\delta) \subseteq \bigcap_{\varepsilon>0} \text{supp}_{2\varepsilon}(\delta) = \bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$.)

For $\eta > 0$, define the period 1 recommendation distribution under $\mu_{\text{product}}(\emptyset)$ by

$$\mu_{\text{product}}(\emptyset)(a_1) = \begin{cases} (1-\eta)\mu(\emptyset)(a_1) & \text{if } a_1 \in \text{supp}(\mu(\emptyset)), \\ \dfrac{\eta}{\left|\prod_{i\in I} \text{supp}(\mu_i(\emptyset))\right| - |\text{supp}(\mu(\emptyset))|} & \begin{array}{c} \text{if } a_1 \in \prod_{i\in I} \text{supp}(\mu_i(\emptyset)) \\ \text{but } a_1 \notin \text{supp}(\mu(\emptyset)), \end{array} \\ 0 & \text{otherwise.} \end{cases}$$

Note that the mediator recommends action a_1 with a positive probability if $a_{i,1}$ is in the support of the marginal distribution of player i's recommendation under μ. In particular, $\text{supp}\left(\mu_{\text{product}}(\emptyset)\right) = \prod_{i\in I} \text{supp}(\mu_i(\emptyset))$.

In subsequent periods $t \geq 2$, if $a_1 \in \text{supp}(\mu(\emptyset))$ was recommended in period 1, then the mediator recommends actions according to $\mu_{\text{product}}(h_m^t) = \mu(h_m^t)$. If $a_1 \in \prod_{i\in I} \text{supp}(\mu_i(\emptyset))$ but $a_1 \notin \text{supp}(\mu(\emptyset))$ was recommended, then the mediator "resets" the history and recommends actions according to μ as if the game started from period 2: formally, $\mu_{\text{product}}(h_m^t) = \mu(\tilde{h}_m^{t-1})$ with $\left((r_1, a_1), \tilde{h}_m^{t-1}\right) = h_m^t$.

As $\eta \to 0$, player i's belief about $a_{-i,1}$ given $a_{i,1} \in \text{supp}(\mu_i(\emptyset))$ under μ_{product} converges to her belief given $a_{i,1}$ in the original equilibrium μ. Since μ_{product} is an ε-sequential equilibrium, this implies that player i's deviation gain in period 1 converges to at most ε as $\eta \to 0$. In addition, if $a_1 \in \text{supp}(\mu(\emptyset))$, then continuation play from period 2 on is the same under μ_{product} and μ. Furthermore, if $a_1 \in \prod_{i \in I} \text{supp}(\mu_i(\emptyset))$ but $a_1 \notin \text{supp}(\mu(\emptyset))$, then continuation play from period 2 on under μ_{product} is the same as play from period 1 on under μ. In sum, for any $\varepsilon' > \varepsilon$, there exists $\eta > 0$ such that player i's deviation gain at any history under μ_{product} is at most ε', so μ_{product} is an ε'-sequential equilibrium strategy. ■

Second, we show that $\text{supp}_\varepsilon^t(\delta)$ is time invariant, and the on-path support and off-path support coincide:

Lemma 11 $\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta) = \bigcap_{\varepsilon>0} \bigcup_t \text{supp}_\varepsilon^t(\delta)$.

Proof. Since $\bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta) \subseteq \bigcap_{\varepsilon>0} \bigcup_t \text{supp}_\varepsilon^t(\delta)$ by definition, we are left to show $\bigcap_{\varepsilon>0} \bigcup_t \text{supp}_\varepsilon^t(\delta) \subseteq \bigcap_{\varepsilon>0} \text{supp}_\varepsilon(\delta)$. Fix an ε-sequential equilibrium strategy μ with a consistent belief system β. As in Lemma 10, it suffices to show that, for each t and $\varepsilon' > \varepsilon$, there exists an ε'-sequential equilibrium strategy $\bar{\mu}$ with consistent belief system $\bar{\beta}$ such that $\text{supp}\left(\bar{\mu}(\emptyset)\right) = \bigcup_{h_m^t} \text{supp}(\mu(h_m^t))$, where the union is taken over all histories h_m^t, including off-path histories.

Fix t and $\varepsilon' > \varepsilon$ arbitrarily. Then, there exists $\eta > 0$ such that if, for each i, τ, h_i^τ, and h_m^τ, a strategy-belief system pair $\left(\bar{\mu}, \bar{\beta}\right)$ satisfies

$$\bar{\mu}(h_m^\tau) = \mu(h_m^\tau) \tag{19}$$

and

$$\bar{\beta}(h_m^\tau | h_i^\tau) \in \left[(1-\eta)^2 \beta(h_m^\tau | h_i^\tau), \frac{1}{(1-\eta)^2} \beta(h_m^\tau | h_i^\tau)\right], \tag{20}$$

that is, if the recommendation schedule is the same as μ and the posterior is close to β, then $\left(\bar{\mu}, \bar{\beta}\right)$ is ε'-sequentially rational. Fix such $\eta > 0$.

Since μ is an ε-sequential equilibrium, there exists a sequence of completely mixed strategies $(\sigma^n)_{n=1}^{\infty}$ and associated belief systems $(\beta^n)_{n=1}^{\infty}$ such that σ^n converges to the obedient strategy profile and β^n converges to β. Consider the following mediator's strategy $\bar{\mu}$. In period 1, the mediator draws a "fictitious" history h_m^t according to $\beta^n(h_m^t)$, the probability of h_m^t given σ^n. Given h_m^t, the mediator draws r according to $\mu(h_m^t)$. In period 1, the mediator sends (h_i^t, r_i) to each player i. Note that, since σ^n is completely mixed, we have $\operatorname{supp}(\bar{\mu}(\emptyset)) = \bigcup_{h_m^t} \operatorname{supp}(\mu(h_m^t))$.

In period $\tau \geq 2$, let $h_m^{2,\tau} = \left(a_1, (r_{\tau'}, a_{\tau'})_{\tau'=2}^{\tau-1}\right)$ with $h_m^{2,2} = a_1$ be the mediator's history at the beginning of period τ except for her recommendation r in period 1. The mediator recommends actions as if the history $h_m^{2,\tau}$ happens after (h_m^t, r) in the original equilibrium μ: that is, she recommends action profiles according to $\mu(h_m^t, r, h_m^{2,\tau})$, where $(h_m^t, r, h_m^{2,\tau})$ denotes the concatenation of h_m^t, r, and $h_m^{2,\tau}$. Similarly, let $h_i^t = (r_{i,s}, a_s)_{s=1}^{t-1}$ be player i's fictitious history, and let $h_i^{2,\tau} = \left(a_1, (r_{i,\tau'}, a_{\tau'})_{\tau'=2}^{\tau-1}\right)$ be her history at the beginning of period τ.

When we view $\mathfrak{h}_m^\tau = (h_m^t, r, h_m^{2,\tau})$ and $\mathfrak{h}_i^\tau = \left(h_i^t, r_i, h_i^{2,\tau}\right)$ as the history (including the fictitious one drawn at the beginning of the game), this new strategy $\bar{\mu}$ satisfies (19). Hence, it suffices to show that we can construct a consistent belief system $\bar{\beta}$ which satisfies (20). To this end, denote player i's belief about the mediator's history $\mathfrak{h}_m^\tau$ when player i's history in period τ is $\mathfrak{h}_i^\tau = \left(h_i^t, r_i, h_i^{2,\tau}\right)$ by $\Pr\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right)$.

We construct the following sequence of perturbed full-support strategies: Each player i plays $\sigma_i^k(h_i^t, r_i)$ in period 1 and then plays $\sigma_i^k(h_i^t, r_i, h_i^{2,\tau}, r_{i,\tau})$ in period $\tau \geq 2$. Recall that $\left(\sigma^k\right)_{k=1}^{\infty}$ is the sequence converging to the original equilibrium σ.

Let $\Pr^{\sigma^n,\sigma^k}\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right)$ be player i's belief about $\mathfrak{h}_m^\tau$ conditional on her history $\mathfrak{h}_i^\tau$, given that the fictitious history h_m^t is drawn from β^n at the beginning of the game and the players take σ^k in the subsequent periods. We have

$$\begin{aligned}
\Pr^{\sigma^n,\sigma^k}\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right) &= \frac{\Pr^{\sigma^k}\left(h_m^{2,\tau} \mid h_m^t, r\right)\Pr^{\mu}\left(r \mid h_m^t\right)\Pr^{\sigma^n}\left(h_m^t\right)}{\Pr^{\sigma^k}\left(h_i^{2,\tau} \mid h_i^t, r_i\right)\Pr^{\mu}\left(r_i \mid h_i^t\right)\Pr^{\sigma^n}\left(h_i^t\right)} \\
&= \frac{\Pr^{\sigma^k}\left(h_m^{2,\tau} \mid h_m^t, r\right)\Pr^{\mu}\left(r \mid h_m^t\right)}{\sum_{\tilde{h}_m^t,\tilde{r}_{-i}} \Pr^{\sigma^k}\left(h_i^{2,\tau} \mid \tilde{h}_m^t, r_i, \tilde{r}_{-i}\right)\Pr^{\mu}\left(r_i, \tilde{r}_{-i} \mid \tilde{h}_m^t\right)\Pr^{\sigma^n}\left(\tilde{h}_m^t \mid h_i^t\right)} \frac{\Pr^{\sigma^n}\left(h_m^t\right)}{\Pr^{\sigma^n}\left(h_i^t\right)}.
\end{aligned}$$

By consistency of (μ, β), for sufficiently large n, we have

$$(1-\eta)\,\beta\left(h_m^t \mid h_i^t\right) \leq \frac{\Pr^{\sigma^n}\left(h_m^t\right)}{\Pr^{\sigma^n}\left(h_i^t\right)} = \beta^n\left(h_m^t \mid h_i^t\right) \leq \frac{1}{1-\eta}\beta\left(h_m^t \mid h_i^t\right)$$

for all h_m^t and h_i^t. Moreover,

$$(1-\eta)\,\beta\left(\tilde{h}_m^t \mid h_i^t\right) \leq \Pr^{\sigma^n}\left(\tilde{h}_m^t \mid h_i^t\right) = \beta^n\left(\tilde{h}_m^t \mid h_i^t\right) \leq \frac{1}{1-\eta}\beta\left(h_m^t \mid h_i^t\right)$$

for each $\tilde{h}_m^t$ and h_i^t.

Fix such a sufficiently large n, and define

$$\bar{\beta}\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right) = \lim_{k\to\infty} \Pr^{\sigma^n,\sigma^k}\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right).$$

Since the belief system β is consistent, we have

$$\lim_{k\to\infty} \Pr^{\sigma^k}\left(h_m^{2,\tau} \mid h_m^t, r\right) = \beta(h_m^{2,\tau} \mid h_m^t, r)$$

Hence, for each $\mathfrak{h}_m^\tau$ and $\mathfrak{h}_i^\tau$, we have

$$\begin{aligned}
\bar{\beta}\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right) &= \frac{\beta\left(h_m^{2,\tau} \mid h_m^t, r\right)\Pr^{\mu}\left(r \mid h_m^t\right)\Pr^{\sigma^n}\left(h_m^t \mid h_i^t\right)}{\sum_{\tilde{h}_m^t,\tilde{r}_{-i}} \beta\left(h_i^{2,\tau} \mid \tilde{h}_m^t, r_i, \tilde{r}_{-i}\right)\Pr^{\mu}\left(r_i, \tilde{r}_{-i} \mid \tilde{h}_m^t\right)\Pr^{\sigma^n}\left(\tilde{h}_m^t \mid h_i^t\right)} \\
&= \left[(1-\eta)^2\,\beta\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right), \frac{1}{(1-\eta)^2}\beta\left(\mathfrak{h}_m^\tau \mid \mathfrak{h}_i^\tau\right)\right].
\end{aligned}$$

That is, $\bar{\beta}$ satisfies (20). ■

Third, we recall that the ε-sequential equilibrium correspondence is upper-hemicontinuous in ε in the product topology (Fudenberg and Levine, 1983). Recall that $\operatorname{supp}_\varepsilon(\delta)$ is the support of on-path period-1 recommendations, while $\bigcup_t \operatorname{supp}_\varepsilon^t(\delta)$ includes off-path recommendations. Hence, by upper-hemicontinuity, we have $\bigcap_{\varepsilon>0}\operatorname{supp}_\varepsilon(\delta) \subseteq \operatorname{supp}(\delta) \subseteq \bigcap_{\varepsilon>0}\bigcup_t \operatorname{supp}_\varepsilon^t(\delta)$. By Lemmas 10 and 11, we have $\bigcap_{\varepsilon>0}\operatorname{supp}_\varepsilon(\delta) = \operatorname{supp}(\delta) = \bigcap_{\varepsilon>0}\bigcup_t \operatorname{supp}_\varepsilon^t(\delta)$, and $\operatorname{supp}(\delta)$ is a product set.

Finally, we prove that $\text{supp}(\delta) = \mathcal{A}^\infty$ whenever $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset$. We first show that $\text{supp}(\delta)$ is a fixed point of the operator S.

Lemma 12 $\text{supp}(\delta) = S(\text{supp}(\delta))$.

Proof. As $S(\bar{A}) \subseteq \bar{A}$ for every product set $\bar{A}$, it suffices to show that $\text{supp}(\delta) \subseteq S(\text{supp}(\delta))$.

Fix $a \in \text{supp}(\delta)$, and let μ be an equilibrium strategy such that $a \in \text{supp}(\mu(\emptyset))$ (which exists by Lemma 11 and upper-hemicontinuity). Let α_{-i} be the conditional distribution of i's opponents' actions in period 1 given recommendation a_i. By incentive compatibility in period 1,

$$(1-\delta) u_i(a_i, \alpha_{-i}) + \delta w_i(a_i|a_i) \geq \max_{\hat{a}_i \in A_i} \{(1-\delta) u_i(\hat{a}_i, \alpha_{-i}) + \delta w_i(\hat{a}_i|a_i)\}, \tag{21}$$

where $w_i(a_i'|a_i)$ is player i's expected continuation payoff from playing a_i' when she is recommended a_i in period 1 under μ. By definition of $\text{supp}(\delta)$, $\alpha_{-i} \in \Delta(\text{supp}(\delta))$. In addition, player i's opponents only play actions in $\text{supp}(\delta)$ on- and off-path. Hence,

$$w_i(a_i|a_i) \leq \max_{a_i \in A_i, \bar{a}_{-i} \in \bar{A}_{-i}} u_i(a_i, \bar{a}_{-i}).$$

In addition, player i's history after she deviates to $\hat{a}_i$ in period 1 is consistent with her opponents' equilibrium strategies, so sequential rationality of player i implies that

$$w_i(\hat{a}_i|a_i) \geq \min_{\hat{\alpha}_{-i} \in \Delta(\bar{A}_{-i})} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$$

Therefore, (21) implies that (10) holds. Hence, $a_i \in S_i(\text{supp}(\delta))$ for all $i \in I$, so $a \in S(\text{supp}(\delta))$. ■

As S is monotone and $\text{supp}(\delta)$ is a fixed point of S, it follows that $\text{supp}(\delta) \subseteq \mathcal{A}^n$ for all n, and hence $\text{supp}(\delta) \subseteq \mathcal{A}^\infty$. The following lemma therefore completes the proof.

Lemma 13 *If* $\text{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u(\mathcal{A}^\infty)\right) \neq \emptyset$, *then* $\mathcal{A}^\infty \subseteq \text{supp}(\delta)$.

Proof. Fix $v \in \mathrm{int}\left(\bigcap_{i\in I} \bar{W}_i \cap u\left(\mathcal{A}^\infty\right)\right) \neq \emptyset$, and let $\mu \in \Delta\left(\mathcal{A}^\infty\right)$ be such that $u\left(\mu\right) = v$ and $\mu\left(r\right) > 0$ for all $r \in \mathcal{A}^\infty$. Let α^*_{-i} be a solution to the problem

$$\min_{\hat{\alpha}_{-i}\in\Delta(\mathcal{A}^\infty)} \max_{a_i\in A_i} u_i\left(a_i, \alpha_{-i}\right).$$

Let α^ε_{-i} be the following full-support (within $\mathcal{A}^\infty$) approximation of α^*_{-i}: $\alpha^\varepsilon_{-i} = (1-\varepsilon)\,\alpha^*_{-i} + \varepsilon \sum_{a_{-i}\in\mathcal{A}^\infty_{-i}} \frac{a_{-i}}{\left|\mathcal{A}^\infty_{-i}\right|}$. Since $v \in \mathrm{int}\left(\bigcap_{i\in I} \bar{W}_i\right)$, there exists $\varepsilon > 0$ such that, for each $i \in I$, we have

$$v_i > \max_{a_i\in A_i} u_i\left(a_i, \alpha^\varepsilon_{-i}\right) + \frac{1-\delta}{\delta} \max_{r\in\mathcal{A}^\infty, a_i\in A_i} \left\{u_i(a_i, r_{-i}) - u_i(r)\right\}. \tag{22}$$

In addition, choose ε small enough such that, for each player i, some best response to α^ε_{-i} is included in $\mathcal{A}^\infty_i$: this is always possible, as every best response to α^*_{-i} is included in $\mathcal{A}^\infty_i$.

We now construct an equilibrium strategy μ^* with $\mathrm{supp}\left(\mu^*\left(\emptyset\right)\right) = \mathcal{A}^\infty$. The construction mirrors the proof of Lemma 1.

The mediator follows an automaton strategy whose state is identical to a subset of players, $J \subseteq I$. If $J = \emptyset$, then the mediator recommends μ. If there exists i with $J = \{i\}$, then the mediator recommends r_{-i} to players $-i$ according to α^ε_{-i}, and recommends some best response to α^ε_{-i} that lies in $\mathcal{A}^\infty_i$ to player i. If $|J| \geq 2$, then for each $i \in J$, the mediator recommends some best response to α^ε_{-i} that lies in $\mathcal{A}^\infty_i$, while she recommends r_{-J} to the other players $-J$ according to $\frac{a_{-J}}{\left|\mathcal{A}^\infty_{-J}\right|}$. The state transits as follows: if the current state is J and players J' deviate, then the state transits to $J \cup J'$.

Player i's strategy is to follow the recommendation $r_{i,t}$ in period t. She believes that the mediator's state is $\emptyset$ 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 within $\mathcal{A}^\infty$, player i's belief is consistent. Given this belief, if player i has deviated, it is optimal for her to always play a static best response to α^ε_{-i}, since the mediator always recommends α^ε_{-i} in state $\{i\}$. Given that deviations by player i are punished in this way, (22) implies that player i has a strict incentive to follow her recommendation r_i at any recommendation profile $r \in \mathcal{A}^\infty$. Hence, she has a strict incentive to follow her recommendation when she believes that r_{-i} is distributed according to $\Pr^{\mu^*}(r_{-i}|r_i)$. ■

9.5 Proof of Proposition 5

Consider the following five-player game with common discount factor $\delta = \frac{2}{3}$. Player 5 has one action and her utility is always equal to $-u_1(a)$. Player 1 has two actions $\{1_1, 2_1\}$ and players 2 and 4 have three actions respectively: $a_i \in \{1_i, 2_i, 3_i\}$ for each $i \in \{2, 4\}$. Finally, player 3 has four actions $a_3 \in \{1_3, 2_3, 3_3, 4_3\}$.

When player 3 takes 1_3, then the payoff matrix for players 1 and 2 is represented as follows, regardless of player 4's actions:

	1_2	2_2	3_2
1_1	$0, 1$	$K, 5$	$K, 1$
2_1	$-2, 0$	$K, 0$	$K, 4$

In addition, player 3's payoff is 0 and player 4's payoff is 0, regardless of (a_1, a_2, a_4). Here, K is a large number.

When player 3 takes 2_3, 3_3, or 4_3, then the payoffs are determined by the following matrix, regardless of (a_1, a_2):

	1_4	2_4	3_4
2_3	$0, 1, 0, 0$	$K, 0, K, K$	$K, 0, K, -K$
3_3	$1, 1, 0, 0$	$K, 0, K, -K$	$K, 0, K, K$
4_3	$1, 1, 0, 0$	$K, 0, K, -K$	$K, 3, K, K$

Note that each player's minmax payoff is equal to zero.

Suppose that the Pareto weight is $\lambda = (0, 0, 0, 0, 1)$: that is, we want to maximize player 5's payoff. We show that player 5's best payoff is strictly less than zero with mediated perfect monitoring , while her best payoff equals zero with some private monitoring structure.

Proposition 7 *There exists K such that* $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ *and* $(v_1, v_2, v_3, v_4, 0) \notin \overline{E_{\text{med}}(\delta)}$ *for all* (v_1, v_2, v_3, v_4)*, and yet there exists a private monitoring structure* p *such that* $(v_1, v_2, v_3, v_4, 0) \in E(p, \delta)$ *for some* (v_1, v_2, v_3, v_4)*.*

9.5.1 Mediated Perfect Monitoring

Since $u_5(a) = -u_1(a)$, it suffices to show that there exists $\eta > 0$ such that, for sufficiently large K, we have

$$(1-\delta)\,\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^{t-1} u_1(a_t)\right] > \eta$$

in every equilibrium with mediated perfect monitoring.

We take K sufficiently large so that $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$. For the sake of contradiction, suppose that such an η does not exist: $\overline{E_{\text{med}}(\delta)}$ includes a payoff profile v with $v_1 = 0$. Since $\overline{E_{\text{med}}(\delta)}$ is equal to the set of (possibly weak) sequential equilibrium payoffs with $\text{int}\left(\bigcap_{i\in I} W_i \cap u(A)\right) \neq \emptyset$ (see the proof of Lemma 6), it suffices to show that the set of (possibly weak) sequential equilibrium payoffs, denoted by $E_{\text{med}}^{\text{weak}}(\delta)$, cannot contain v with $v_1 = 0$.

Note that $(v_1, v_2) = (0, 1)$ is not in $E_{\text{med}}^{\text{weak}}(\delta)$ for any (v_2, v_3, v_5). Since $(0, 1, 0, 0, 0)$ is an extreme point in the feasible set, we have to recommend $(1_1, 1_2, 1_3, a_4)$ or $(a_1, a_2, 2_3, 1_4)$ with probability one. However, $(1_1, 1_2, 1_3, a_4)$ requires a high continuation payoff for player 2 since, if she deviates to 4_2 from 1_2, her gain is 4. On the other hand, $(a_1, a_2, 2_3, 1_4)$ requires a high continuation payoff for player 4 since, if she deviates to 1_4 from 2_4, her gain is K. By feasibility, we have to guarantee a positive continuation payoff for player 1. Hence, $(v_1, v_2) = (0, 1)$ is not in $E_{\text{med}}^{\text{weak}}(\delta)$.

Since $E_{\text{med}}^{\text{weak}}(\delta)$ is closed (by the usual argument with product topology) and $(0, 1)$ is an extreme point, there exists $\bar{\eta} > 0$ such that, given $v_1 < \bar{\eta}$, we should have $v_2 < 1 - \bar{\eta}$.

Let $\Pr(a)$ be the probability that the players take an action profile a in period 1 in the equilibrium. Player 1's payoff is bounded from below by

$$(1-\delta)\left\{\Pr(2_3)\Pr(2_4 \text{ or } 3_4 \mid 2_3)K + \Pr(3_3) + \Pr(4_3)\right\},$$

since from period 2 on she can guarantee the payoff of 0 by taking 1_1. Hence, we need to have $\Pr(3_3) = \Pr(4_3) = 0$. Given that $\Pr(3_3) = \Pr(4_3) = 0$, we should have $\Pr(2_3) = 0$ as well. To see why, suppose otherwise. Then, we should have $\Pr(2_4 \text{ or } 3_4 \mid 3_3) = 0$, that is, $\Pr(1_4 \mid 2_3) = 1$. Hence, player 4's deviation gain from 1_4 to 2_4 is no less than

$$K \Pr(2_3 \mid 1_4) \geq K \Pr(1_4 \mid 2_3) \Pr(2_3) = K \Pr(2_3).$$

Therefore, so that player 4 does not deviate, we need to guarantee the payoff of $\frac{1-\delta}{\delta} K \Pr(2_3)$. By feasibility, we need to guarantee a positive payoff to player 1. This is a contradiction. In total, we should have $\Pr(2_3) = \Pr(3_3) = \Pr(4_3) = 0$.

Given $\Pr(2_3) = \Pr(3_3) = \Pr(4_3) = 0$, let us focus on the payoff matrix given $a_3 = 1_3$:

	1_2	2_2	3_2
1_1	$0, 1$	$K, 5$	$K, 1$
2_1	$-2, 0$	$K, 0$	$K, 4$

Player 1's payoff is bounded from below by $(1 - \delta) \{1 - \Pr(1_2)\} K$. Hence, we should have $\Pr(1_2) = 1$

Given $a_2 = 1_2$, player 2's deviation gain when she takes 4_2 is

$$\Pr(1_1 \mid 1_2)4 - \Pr(2_1 \mid 1_2)0,$$

and her deviation gain when she takes 5_2 is

$$-\Pr(1_1 \mid 1_2)0 + \Pr(2_1 \mid 1_2)4.$$

On the other hand, given $a_1 = 1_1$, player 1's deviation gain is zero. Given $a_1 = 2_1$, her deviation gain is

$$\Pr(1_2 \mid 1_1) 2.$$

Let $p = \Pr(1_1)$. Since $\Pr(1_2) = 1$, $\Pr(1_1 \mid 1_2) = p$ and $\Pr(2_1 \mid 1_2) = 1 - p$. The smallest continuation payoff w_2 that we need to guarantee player 2 is

$$\frac{1}{2}\max\{4p, 4 - 4p\} = \max\{2p, 2 - 2p\},$$

while the continuation payoff that we need to guarantee player 1 is

$$\begin{cases} 0 & \text{after } a_1 = 1_1 \text{ is recommended,} \\ 1 & \text{after } a_1 = 2_1 \text{ is recommended.} \end{cases}$$

Hence, the continuation payoffs after $a_1 \in \{1_1, 2_1\}$, denoted by $w(1_1)$ and $w(2_1)$, should satisfy the following:

$$\begin{cases} pw_2(1_1) + (1-p)\,w_2(2_1) \geq \max\{2p, 2-2p\}, \\ w_1(1_1) \geq 0, \\ w_1(2_1) \geq 1. \end{cases}$$

Note that player 1's payoff is no less than

$$p\delta w_1(1) + (1-p)\{(1-\delta)(-2) + \delta w_1(2_1)\}.$$

Hence, $w_1(1)$ needs to be sufficiently close to 0 or p is close to zero. However, if p is close to zero, then we should have $w_1(2_1)$ close to 1 and $w_2(2_1)$ close to 2. By feasibility, this is impossible. Therefore, $w_1(1)$ should be close to 0. The symmetric argument shows that $w_1(2_1)$ is sufficiently close to 1.

In addition, recall that there exists $\bar{\eta} > 0$ such that, given $v_1 < \bar{\eta}$, we should have $v_2 < 1 - \bar{\eta}$. Hence, after $a_1 = 1_1$, we should have $w_2(1_1) < 1 - \bar{\eta}$. Hence, the first constraint implies

$$p(1-\bar{\eta}) + (1-p)\,w_2(2_1) \geq \max\{2p, 2-2p\}.$$

For $p \leq \frac{1}{2}$, this is equivalent to

$$p\,(1-\bar{\eta})+(1-p)\,w_2(2_1) \geq 2-2p,$$

or

$$w_2(2_1) \geq \frac{1}{1-p}\,(2-3p+p\bar{\eta})\,.$$

For sufficiently small $\bar{\eta}$, the left hand side is decreasing in p. Hence, we should have

$$w_2(2_1) \geq 1+\bar{\eta}.$$

By feasibility, for sufficiently large K, this means that $w_2(1_1)$ is sufficiently large, which is a contradiction. A symmetric argument shows that $p > \frac{1}{2}$ leads us to a contradiction as well.

9.5.2 Private Monitoring

With the following private monitoring, $(0,1,0,0,0) \in E(\delta,p)$: Players 1, 2, 3, and 5 observe the other players' actions and their own payoffs perfectly. Player 4 cannot observe anything but her own payoff.

On the equilibrium path, in period 1, players take $(1_1, 1_2, 1_3, 1_4)$ and $(2_1, 1_2, 1_3, 1_4)$ with probabilities $\frac{1}{2}$ and $\frac{1}{2}$, respectively (that is, player 1 mixes 1_1 and 2_1 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ and the other takes 1_i). From period 2 on, player 3 takes 2_3 if player 1 took 1_1 in period 1, and she takes 3_3 if player 1 took 2_1. Players 1, 2, and 4 take $(1_1, 1_2, 1_4)$ with probability one.

Off the equilibrium path, all players' deviation are ignored, except for player 1's deviation in period 1 and player 4's deviation in period $t \geq 2$. If player 1 deviates in period 1 (or player 4 deviates in period $t \geq 2$), then in each period $\tau \geq 2$ (or in each period $\tau \geq t+1$,) players 1 and 2 takes 1_1 and 1_2 with probability one, player 3 mixes 2_3 and 3_3 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ respectively, *i.i.d.* across periods, and player 4 mixes 2_4 and 3_4 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ respectively, *i.i.d.* across periods. Note that this is a static Nash equilibrium.

Let us verify the players' incentives: Player 3 takes a static best response. Players 1 and 2 take a static best response from period 2 on, on and off the equilibrium path. Consider player 4's incentive from period 2 on. In period 1, player 4's payoff is constant at 0 regardless of her actions. Hence, player 4 in period 2 believes that player 1 mixed 1_1 and 2_1 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ in period 1, and so player 3 takes 2_3 and 3_3 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ respectively in period 2. Since player 4's payoff is constant as long as she takes 1_4, on the equilibrium path, player 4 in period t believes that player 3 takes 2_3 and 3_3 with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ respectively in period t. Hence, instantaneous deviation gain is zero while she will be minmaxed from the next period once she deviates. Hence, player 4 has the incentive to follow the equilibrium from period 2 on. Since player 4's deviation in period 1 does not change player 4's payoff or information, player 4 has the incentive to follow the equilibrium in period 1 as well.

Hence, we are left to verify players 1 and 2's incentives in period 1. For player 1, recalling that $\delta = \frac{2}{3}$, we have

$$\begin{aligned}
\text{payoff of taking } 1_1 &= (1-\delta)\,0 + \delta \times \underbrace{0}_{\text{after } 1_1 \text{, player 3 takes } 2_3 \text{ with probability one}} \\
&= (1-\delta)\,(-2) + \delta \times \underbrace{1}_{\text{after } 2_1 \text{, player 3 takes } 3_3 \text{ with probability one}} \\
&= \text{payoff of taking } 2_1.
\end{aligned}$$

Hence, the equilibrium is incentive compatible for player 1. For player 2,

$$\begin{aligned}
\text{payoff of taking } 1_2 &= (1-\delta)\left(\frac{1}{2}\times 1 + \frac{1}{2}\times 0\right) + \delta \times \underbrace{1}_{\text{whether player 3 takes } 2_3 \text{ or } 3_3 \text{, player 2 obtains 1}} \\
&\geq (1-\delta)\left(\frac{1}{2}\times 5\right) + \delta \times 0 \\
&= \text{payoff of taking } 2_2 \\
&= (1-\delta)\left(\frac{1}{2}\times 1 + \frac{1}{2}\times 4\right) + \delta \times 0 \\
&= \text{payoff of taking } 3_2.
\end{aligned}$$

Hence, the equilibrium is incentive compatible for player 2.

9.6 Proof of Corollary 1

Since $u(A)$ is convex and $v > \underline{u}$, there exists $v' \in u(A)$ and $\varepsilon > 0$ such than $v - 2\varepsilon > v' > \underline{u}$. Taking $\bar{\delta}_1 = \max_{i \in I} \frac{d_i}{v'_i - \underline{u}_i + d_i}$ and applying Proposition 6 yields $v' \in \overline{E_{\text{med}}(\delta)}$ for all $\delta > \bar{\delta}_1$. Thus, for every $\delta > \bar{\delta}_1$, there exists $v(\delta) \in E_{\text{med}}(\delta)$ such that $v - \varepsilon > v(\delta) > \underline{u}$.

Let $\bar{\delta}_2 = \max_{i \in I} \frac{d_i}{\varepsilon + d_i}$ and let $\bar{\delta} = \max\left\{\bar{\delta}_1, \bar{\delta}_2\right\}$. Then, for every $\delta > \bar{\delta}$, the following is an equilibrium: On-path, the mediator recommends a correlated action profile that attains v. If anyone deviates, play transitions to a sequential equilibrium yielding payoff $v(\delta)$. Such an equilibrium exists as $v(\delta) \in E_{\text{med}}(\delta)$ for $\delta > \bar{\delta} \geq \bar{\delta}_1$. Finally, since this punishment results in a loss of continuation payoff of at least ε for each player, the mediator's on-path recommendations are incentive compatible for $\delta > \bar{\delta} \geq \bar{\delta}_2$.

References

[1] Abreu, D., P.K. Dutta, and L. Smith (1994), "The Folk Theorem for Repeated Games: a NEU Condition," *Econometrica*, 62, 939-948.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1063.

[3] Amarante, M. (2003), "Recursive Structure and Equilibria in Games with Private Monitoring," *Economic Theory*, 22, 353-374.

[4] Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32, 428-465.

[5] Awaya, Y. and V. Krishna (2014), "On Tacit versus Explicit Collusion," *mimeo.*

[6] Ben-Porath, E. and M. Kahneman (1996), "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory*, 70, 281-297.

[7] Bergemann, D. and S. Morris (2013), "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251-1308.

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

[9] Compte, O. (1998), "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597-626.

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

[11] Forges, F. (2009), "Correlated Equilibrium and Communication in Games," *Encyclopedia of Complexity and Systems Science*, edited by R. Meyers, Springer: New York.

[12] Forges, F., J.-F. Mertens, and A. Neyman (1986), "A Counterexample to the Folk Theorem with Discounting," *Economics Letters*, 20, 7.

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

[14] Fudenberg, D. and D. Levine (1983), "Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games," *Journal of Economic Theory*, 31, 251-268.

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

[16] Green, E.J. and R.H. Porter (1984), "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87-100.

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

[18] Kandori, M. (2002), "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1-15.

[19] Kandori, M. and H. Matsushima (1998), "Private Observation, Communication and Collusion," *Econometrica*, 66, 627-652.

[20] Kandori, M. and I. Obara (2010), "Towards a Belief-Based Theory of Repeated Games with Private Monitoring: An Application of POMDP," *mimeo*.

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

[22] MacLeod, B.W. (2003), "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 93, 216-240.

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

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

[25] Miyahara, Y. and T. Sekiguchi (2013), "Finitely Repeated Games with Monitoring Options," *Journal of Economic Theory*, 148, 1929-1952.

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

[27] Phelan, C., and A. Skrzypacz (2012), "Beliefs and Private Monitoring," *Review of Economic Studies*, 79, 1637-1660.

[28] Phelan, C. and E. Stacchetti (2001), "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69, 1491-1518.

[29] Rahman, D. (2012), "But Who Will Monitor the Monitor?," *American Economic Review*, 102, 2767-2797.

[30] Rahman, D. (2014), "The Power of Communication," *American Economic Review*, 104, 3737-51.

[31] Rahman, D. and I. Obara (2010), "Mediated Partnerships," *Econometrica*, 78, 285-308.

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

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

[34] Sugaya, T. and A. Wolitzky (2014a), "Characterization of Strict Equilibrium Payoffs in Repeated Games with Mediated Perfect Monitoring," *mimeo.*

[35] Sugaya, T. and A. Wolitzky (2014b), "Perfect Versus Imperfect Monitoring in Repeated Games," *mimeo.*

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

Online Appendix: Proof of Proposition 4

The following preliminary result is useful: We show that a best response to the other players' strategy is included in the support:

Lemma 14 *For each $i \in I$, if $\alpha_{-i} \in \Delta\left(\text{supp}_{-i}(\delta)\right)$, then a static best response to α_{-i}, denoted by $BR_i(\alpha_{-i})$, is included in* $\text{supp}_i(\delta)$.

Since $\text{supp}(\delta)$ has a product structure by Lemma 10, this lemma implies that the support of (BR_i, α_{-i}) is included in $\text{supp}(\delta)$.

Proof. Since the ε-sequential equilibrium correspondence is upper-hemicontinuous in ε, the proof of Lemma 11 implies that there exists an equilibrium μ^{full} such that $\mu^{\text{full}}(\emptyset) = \text{supp}(\delta)$.

Again, by the upper-hemicontinuity, it suffices to show that, if $\alpha_{-i} \in \Delta\left(\text{supp}_{-i}(\delta)\right)$, then a static best response $BR_i(\alpha_{-i})$ is included in $\text{supp}_{i,\varepsilon}(\delta)$ for each $\varepsilon > 0$. To this end, for each $\alpha_{-i} \in \Delta\left(\text{supp}_{-i}(\delta)\right)$ and $BR_i(\alpha_{-i})$, consider the following recommendation μ: In period 1, the mediator draws two states $\omega = 0$ and $\omega = 1$ with probabilities $1-\eta$ and η, respectively. In state $\omega = 0$, the mediator recommends actions according to μ^{full}. In state $\omega = 1$, the mediator recommends $(BR_i(\alpha_{-i}), \alpha_{-i})$ in period 1, and then from period 2 on, the mediator "resets" the history and recommends actions according to μ^{full}: formally, $\mu(h_m^t) = \mu^{\text{full}}(\tilde{h}_m^{t-1})$ with $\left((r_1, a_1), \tilde{h}_m^{t-1}\right) = h_m^t$.

For each $\varepsilon > 0$, for sufficiently small $\eta > 0$, this is an ε-sequential equilibrium for the following reason: For each $\rho > 0$, for sufficiently small $\eta > 0$, since μ^{full} has full support, players $-i$ in period 1 believe that the state is $\omega = 0$ with probability no less than $1-\rho$. Hence, the deviation gain is of order ρ. From period 2 on, following the recommendation is optimal given ω. On the other hand, for player i, it is optimal to follow the recommendation conditional on ω in each period. ■

Given this preliminary result, let us prove Proposition 4. Since we focus on the case with two players, whenever we say players i and j, we assume they are different: $i \neq j$.

The monitoring structure p constructed below will involve perfect monitoring of actions together with a particular ex ante correlating device. As the mediator can always do any subsequent talking on the players' behalf, it is immediate that $E_{\text{talk}}(\delta, p) \subseteq E_{\text{med}}(\delta)$. We must show that any payoff vector $v \in E_{\text{med}}(\delta)$ is also in $E_{\text{talk}}(\delta, p)$. Let μ^v be an equilibrium strategy with a mediator that attains v.

The following lemma allows us to focus on the recommendation schedule with uniform full support:

Lemma 15 *For each $\varepsilon > 0$, there exist μ and $\eta > 0$ such that μ is an ε-sequential equilibrium,*

$$\mu(h_m^t)(a) > \eta \text{ for each } h_m^t \text{ and } a \in \text{supp}(\delta), \tag{23}$$

and $\left\|\mathbb{E}^\mu\left[(1-\delta)\sum_t u(a_t)\right] - v\right\| < \varepsilon$.

Since ε is arbitrary and the ε-sequential equilibrium correspondence is upper-hemicontinuous in ε, it suffices to show that, for each $\varepsilon > 0$, we can replicate such μ as an ε'-sequential equilibrium of the game with the initial correlation and cheap talk for each $\varepsilon' > \varepsilon$.

Proof. From μ^v, we create a new equilibrium which has full support within $\text{supp}(\delta)$ in the initial period as follows. Since the ε-sequential equilibrium correspondence is upper-hemicontinuous in ε, the proof of Lemma 11 implies that there exists an equilibrium μ^{full} such that $\mu^{\text{full}}(\emptyset) = \text{supp}(\delta)$. Given $\rho > 0$, suppose that the mediator has two possible states $\omega \in \{0, 1\}$. The state is drawn at the beginning of the game and fixed for the rest of the game: She draws $\omega = 0$ with probability $1 - \rho$ and draws $\omega = 1$ with probability ρ. In state $\omega = 0$, the mediator recommends action profiles according to μ^v while she recommends actions profiles according to μ^{full} in state $\omega = 1$. Let $\bar{\mu}$ denote such a recommendation schedule. Note that $\bar{\mu}$ is an equilibrium.

From $\bar{\mu}$, we create a new ε-sequential equilibrium which has full support within $\text{supp}(\delta)$ every period as follows. In period 1, the mediator starts recommendation according to $\bar{\mu}$. After each period, with probability $\rho > 0$, the mediator "resets" the history and starts over with $\bar{\mu}$. Let μ denote this recommendation schedule. We can see μ as the same recommendation schedule as $\bar{\mu}$ with the discount factor δ being replaced with $(1 - \rho)\,\delta$. Hence, for each $\varepsilon > 0$, for sufficiently small $\rho > 0$, μ is an ε-sequential equilibrium.

In addition, since $\mu(h_m^t)(a) \geq \rho\bar{\mu}(\emptyset)(a)$ and $\bar{\mu}$ has full support within $\text{supp}(\delta)$, with $\eta \equiv \rho \min_a \bar{\mu}(\emptyset)(a)$, we have (23). Moreover, with a sufficiently small ρ, we have $\|\mathbb{E}^{\mu}\left[(1-\delta)\sum_t u(a_t)\right] - v\| < \varepsilon$. ■

We now replicate ε-sequential equilibrium μ with (23) as an ε'-sequential equilibrium of the game with the initial correlation and cheap talk for each $\varepsilon' > \varepsilon$.

Let α_j^* be the minmax strategy of player j with support restricted to $\text{supp}(\delta)$:

$$\min_{\alpha_j \in \Delta_j(\text{supp}(\delta))} \max_{a_i \in A_i} u_i(a_i, \alpha_j).$$

Let α_j^{full} be a full support perturbation of α_j^*, and let BR_i be a best response against α_j^{full}. Note that the support of $\left(BR_i, \alpha_j^{\text{full}}\right)$ is included in $\text{supp}(\delta)$.[18]

Given $\left(BR_i, \alpha_j^{\text{full}}\right)$, it is useful to see the recommendation schedule with μ after on-path history $a^t = (a_\tau)_{\tau=1}^{t-1}$ as follows:

$$\mu(a^t) = \sum_{i \in I} p_i \left(BR_i, \alpha_j^{\text{full}}\right) + (1 - p_1 - p_2)\,\tilde{\mu}(a^t).$$

That is, with probability p_i, the recommendation is $\left(BR_i, \alpha_j^{\text{full}}\right)$, and with probability $1 - p_1 - p_2$, the mediator draws the recommendation so that the total recommendation is the same as $\mu(a^t)$:

$$\tilde{\mu}(a^t)\,(r) = \frac{\mu(a^t)(r) - \sum_{i=1}^{2} p_i 1_{\{BR_i = r_i\}} \alpha_j^{\text{full}}(r_j)}{1 - p_1 - p_2}.$$

[18] This follows from Lemma 14 if the best response to α_j^{full} is unique. Otherwise, we take BR_i so that it is upper-hemicontinuous with respect to the perturbation of α_j^*.

Since μ has uniform full support (23), we can take $p_i > 0$ for each i and $\tilde{\mu}(a^t)(r) > 0$ for each $r \in \text{supp}(\delta)$. Moreover, as $p_1, p_2 \to 0$, when player i receives $r_{i,t} \in \text{supp}_i(\delta)$ after a^t, she believes that the distribution of r_j is $\tilde{\mu}(a^t)(r_j|r_i)$.

Given this definition, we construct a recommendation schedule which replicates μ with initial correlation and cheap talk.

Intuitively, the mediator makes a^t-contingent recommendations for each possible $a^t = (a_\tau)_{\tau=1}^{t-1}$. In particular, we define a state ω_t, an a^t-contingent recommendation $r_t(a^t)$, and another a^t-contingent recommendation $r_{i,t}^{\text{punish}}(a^t)$ for each player i as follows: In the initial period $t = 1$, the mediator draws $\omega_t \in \{0, 1, 2\}$ with the following probabilities: $\omega_t = 0$, 1, and 2 with probability $1 - p_1 - p_2$, p_1, and p_2, respectively. In addition, she draws $r_t(a^t)$ according to $\tilde{\mu}(a^t)$ while she draws $r_{i,t}^{\text{punish}}(a^t)$ according to α_i^{full} (that is, player i punishes player j).

From period $t \geq 2$, if there exists a unique player i who has deviated in the previous periods, then $\omega_t = j$ forever. Otherwise, $\omega_t = 0$, 1, and 2 with probability $1-p_1-p_2$, p_1, and p_2, respectively. In addition, she draws $r_t(a^t)$ according to $\tilde{\mu}(a^t)$ while she draws $r_{i,t}^{\text{punish}}(a^t)$ according to α_i^{full}. Formally, we say that player i has deviated if there exist $\tau \leq t-1$ and a unique i such that (i) for each $\tau' < \tau$, we have $a_{\tau'} = r_{\tau'}(a^{\tau'})$ if $\omega_{\tau'} = 0$ and $a_{n,\tau'} = r_{n,\tau'}^{\text{punish}}(a^{\tau'})$ if $\omega_{\tau'} = n \in \{1, 2\}$ (no deviation happened until period $\tau - 1$) and (ii) $a_{i,\tau} \neq r_{i,\tau}(a^\tau)$ with $\omega_\tau = 0$ or $a_{i,\tau} \neq r_{i,\tau}^{\text{punish}}(a^\tau)$ with $\omega_\tau = i$ (player i deviates in period τ).

Intuitively, if $\omega_t = 0$, then $r_t(a^t)$ is the instruction for the players while if $\omega_t = i \in \{1, 2\}$, then the mediator recommends $r_{i,t}^{\text{punish}}(a^t)$ to player i to punish player j. On the other hand, player j takes $r_{j,t}(a^t)$ given $\omega_t = i$, rather than BR_j on equilibrium path. This is ε'-sequentially rational for the following reason: given $\omega_t \neq j$, player j believes that $\omega_t = 0$ with a high probability for sufficiently small $p_1, p_2 > 0$ since $\mu(a^t)$ satisfies the uniform full support (23).

We now construct a method for the mediator to send the information about

$$\left(\left(\omega_t, r_t\left(a^t\right), r_{1,t}^{\text{punish}}\left(a^t\right), r_{2,t}^{\text{punish}}\left(a^t\right)\right)_{a^t \in A^{t-1}}\right)_{t=1}^{\infty}$$

in the initial period. The key features to establish is that (i) player i does not know the instruction for the other players, that (ii) before player i reaches period t, player i does not know her own recommendations for periods $\tau \geq t$ (otherwise, player i would obtain more information than the original equilibrium and so might want to deviate), and that (iii) no player wants to deviate (in particular, if player i deviates in actions or cheap talk, then the state will be $\omega_t = j$ and she will be minmaxed for a sufficiently long time).

To this end, without loss, let $A_i = \{1_i, ..., n_i\}$ with $n_i = |A_i|$ be player i's action set. For each t and a^t, the mediator draws a random variables $p_j^i(a^t)$ *i.i.d.* uniformly from $\{1, ..., Nn_in_j\}$ with a large N to be determined. Define

$$r_{i,t}^i(a^t) = r_{i,t}(a^t) + p_j^i(a^t) \pmod{Nn_in_j}. \tag{24}$$

Intuitively, $r^i_{i,t}(a^t)$ is the "encoded instruction" of $r_{i,t}(a^t)$, and to obtain $r_{i,t}(a^t)$ from $r^i_{i,t}(a^t)$, player i needs to know $p^i_j(a^t)$. The mediator gives $p^i_j(a^t)$ to player j so that player i does not know $r_{i,t}(a^t)$ until period t. In this proof, in general, the subscript of the variable represents who owns the information and the superscript represents about whose recommendation the variable is about. In period t, player j sends $p^i_j(a^t)$ by cheap talk.

On the other hand, the mediator directly tells $r^{\text{punish}}_{i,t}(a^t)$ to player i. For sufficiently small $p_1, p_2 > 0$, since $\mu(a^t)$ satisfies the uniform full support (23), until period t comes, player i believes that $\omega_t = i$ with a small probability. Hence, knowing $r^{\text{punish}}_{i,t}(a^t)$ does not change player i's incentive much (and we are considering ε'-sequential rationality).

In order to incentivize player i to tell the truth, the equilibrium should embed the mechanism so that if player i tells a lie, then she will be punished. Intuitively, this will be done as follows: imagine that player j receives N "boxes" $n \in \{1, ..., N\}$ for each period t. One box is unlocked and includes $r^i_{i,t}(a^t)$, while all the others are locked. Let $m^j_{j,t}(a^t)$ be the index of the unlocked box for period t and history a^t. Only player j knows $m^j_{j,t}(a^t)$.

Player i, on the other hand, receives N keys (in the formal definition, $\left(p^i_{j,t}(n, a^t)\right)_{n=1}^N$). For $n = m^j_{j,t}(a^t)$ (the key corresponding to the empty box), $p^i_{j,t}(n, a^t)$ corresponds to $p^i_j(a^t) \in \{1, ..., Nn_in_j\}$. On the other hand, for other $n \neq m^j_{j,t}(a^t)$, $p^i_{j,t}(n, a^t)$ corresponds to the passcode (included in $\{1, ..., Nn_in_j\}$). The correct passcode $p^i_{j,t}(n, a^t)$ opens the locked box n if and only if $\omega_t \neq j$.

Player i is asked to send $\left(p^i_{j,t}(n, a^t)\right)_{n=1}^N$. Suppose that player i tells a lie. Since she does not know $m^j_{j,t}(a^t)$, most likely she sends a wrong passcode for some $n \neq m^j_{j,t}(a^t)$. Since we make sure that the wrong passcode cannot open the box n, player j will believe that $\omega_t = j$ (and so player i will be minmaxed). Hence, if player i tells a lie, then she will be punished for the current period.

So that she will be punished for a sufficiently long time, we need to add another complication. The information that player i has, $\left(p^i_{j,t}(n, a^t)\right)_{n=1}^N$, is not exactly equal to the passcode, but this passcode is encoded by the mediator. The key to decode $\left(p^i_{j,t}(n, a^t)\right)_{n=1}^N$ in order to get the true passcode is given to player j. In addition, if the current state is $\omega_t = j$, then player i also receives the key to decode $\left(p^i_{j,t+1}(n, a^{t+1})\right)_{n=1}^N$ for the next period.

Suppose that player i tells a lie in period t. As explained above, it will create $\omega_t = j$ with a high probability even if $\omega_t = 0$. If the true state is $\omega_t = 0$, then player i needs player j's key to decode $\left(p^i_{j,t+1}(n, a^{t+1})\right)_{n=1}^N$ in order to send the correct passcode for the boxes in the next period. However, player j, who believes that $\omega_t = j$, believes that player i has the key. Hence, player j does not send the key if she believes that $\omega_t = j$. Hence, player i cannot send the correct passcode for long periods with a high probability, and will be minmaxed.

Let us formalize the above construction: The mediator draws $m^j_{j,t}(a^t)$ from the uniform distribution over $\{1, ..., N\}$ *i.i.d.* for each player j, each period t, and each a^t. Moreover, the mediator draws four independent random variables $q^j_{j,t}(n, a^t)$, $q^j_{j,t}(n, \omega_t = j, a^t)$, $q^j_{j,t}(n, \omega_t \neq j, a^t)$, and $q^i_{j,t}(n, \omega_t \neq j, a^t)$ from the uniform distribution over $\{1, ..., Nn_in_j\}$ *i.i.d.* for each $i \in \{1, 2\}$, each $n \in \{1, ..., N\}$, each period t, and each a^t.

$q^j_{j,t}(n,a^t)$ corresponds to the boxes. $q^j_{j,t}(n,\omega_t = j, a^t)$ is the key to decode $\left(p^i_{j,t+1}(n,a^{t+1})\right)_{n=1}^N$, which is shared with players i and j given $\omega_t = j$. On the other hand, if $\omega_t \neq j$, then player j has the true key $q^j_{j,t}(n,\omega_t \neq j, a^t)$ while player i receives a dummy key $q^i_{j,t}(n,\omega_t = j, a^t)$.

In addition, the mediator draws $\left\{q^i_{j,t}(n,a^t)\right\}_{n=1}^{Nn_in_j}$ uniformly with the constraint

$$q^i_{j,t}(n,a^t) \neq q^j_{j,t}(n,a^t) \text{ for at least one } n \in \{1,...,Nn_in_j\} \text{ with } n \neq m^j_{j,t}(a^t) \tag{25}$$

$i.i.d.$ for each $i \in \{1,2\}$, each period t, and each a^t. This $q^i_{j,t}(n,a^t)$ is the passcode that will be used if $\omega_t = j$, which cannot open at least one box n.

Given $q^j_{j,t}(n,\omega_t = j,a^t)$, $q^j_{j,t}(n,\omega_t \neq j,a^t)$, and $q^i_{j,t}(n,\omega_t \neq j,a^t)$, we define

$$q^{j,\omega}_{j,t}(n,a^t) = \begin{cases} q^j_{j,t}(n,\omega_t \neq j,a^t) & \text{if } \omega_t \neq j, \\ q^j_{j,t}(n,\omega_t = j,a^t) & \text{if } \omega_t = j \end{cases} \tag{26}$$

while

$$q^{i,\omega}_{j,t}(n,a^t) = \begin{cases} q^i_{j,t}(n,\omega_t \neq j,a^t) & \text{if } \omega_t \neq j, \\ q^j_{j,t}(n,\omega_t = j,a^t) & \text{if } \omega_t = j. \end{cases} \tag{27}$$

Player j observes $q^{j,\omega}_{j,t}(n,a^t)$ while player i observes $q^{i,\omega}_{j,t}(n,a^t)$. That is, if and only if $\omega_t = j$, players i and j share the same key to decode $\left(p^i_{j,t}(n,a^t)\right)_{n=1}^N$:

$$q^{j,\omega}_{j,t}(n,a^t) = q^{i,\omega}_{j,t}(n,a^t). \tag{28}$$

On the other hand, for each period t and a^t, we define $\left(p^i_{j,t}(n,a^t)\right)_{n=1}^N$ as follows: In period 1,

$$p^i_{j,1}(n,\emptyset) = \begin{cases} p^i_{j,1}(a^t) & \text{if } n = m_{j,t}(a^t) \text{ and } \omega_t \neq j, \\ q^j_{j,t}(n,a^t) & \text{if } n \neq m_{j,t}(a^t) \text{ and } \omega_t \neq j, \\ q^i_{j,t}(n,a^t) & \text{if } \omega_t = j. \end{cases}$$

That is, if $\omega_t \neq j$, then the passcodes for the boxes $n \neq m^j_{j,t}(a^t)$ open the boxes (that is, $p^i_{j,1}(n,\emptyset) = q^j_{j,t}(n,a^t)$ if $n \neq m^j_{j,t}(a^t)$ and $\omega_t \neq j$). On the other hand, if $\omega_t = j$, then $q^i_{j,t}(n,a^t) \neq q^j_{j,t}(n,a^t)$ by (25) and cannot open at least one box.

On the other hand, in period $t \geq 2$, then we define

$$p^i_{j,t}(n,a^t) = \begin{cases} p^j_i(a^t) + q^{j,\omega}_{j,t-1}(n,a^{t-1}) & \text{if } n = m_{j,t}(a^t) \text{ and } \omega_t \neq j, \\ q^j_{j,t}(n,a^t) + q^{j,\omega}_{j,t-1}(n,a^{t-1}) & \text{if } n \neq m_{j,t}(a^t) \text{ and } \omega_t \neq j, \\ q^i_{j,t}(n,a^t) + q^{j,\omega}_{j,t-1}(n,a^{t-1}) & \text{if } \omega_t = j. \end{cases} \tag{29}$$

Intuitively, in period $t+1$, player i knows $p^i_{j,t+1}(n,a^{t+1})$ and is asked to send $p^i_{j,t+1}(n,a^{t+1}) - q^{j,\omega}_{j,t}(n,a^t)$ to player j. As mentioned, $p^i_{j,t+1}(n,a^{t+1})$ is the encoding of player i's passcode. By (26) and (27), player j in period t believes that player i needs to know the key to decode $p^i_{j,t+1}(n,a^{t+1})$, $q^{j,\omega}_{j,t}(n,a^t)$, in period $t+1$ only if $\omega_t \neq j$. Hence, if and only if player j

interprets $\omega_t \neq j$, player j sends $q_{j,t}^{j,\omega}(n, a^t)$ to player i at the end of period t. Therefore, if player i tells a lie and creates a situation where player j cannot open a box and interprets $\omega_t = j$ in period t, then player i cannot receive the information about $q_{j,t}^{j,\omega}(n, a^t)$. This is costly since player i cannot decode $p_{j,\tau}^{i}(n, a^\tau)$ and player j will interpret $\omega_\tau = j$ with a high probability in the subsequent periods $\tau = t+1, \ldots$.

Let us formalize the above argument: From $q_{j,t}^{j}(n, a^t)$ and $q_{j,t}^{i}(n, a^t)$, we define $Q_{j,t}(n, a^t)$ recursively as follows: For the initial period,

$$Q_{j,1}^{j}(n, \emptyset) = \begin{cases} 0 & \text{if } n = m_{j,1}^{j}(a^t), \\ q_{j}^{j}(n, \emptyset) & \text{if } n \neq m_{j,1}^{j}(a^t). \end{cases}$$

For the latter periods, we recursively define

$$\begin{aligned} & Q_{j,t}^{j}(n, a^t) \\ = & \begin{cases} Q_{j,t}(n, a^t) & \text{if } n = m_{j,t}^{j}(a^t) \text{ and } \omega_{t-1} \neq j, \\ Q_{j,t}(n, a^t) + q_{j,t}^{j}(n, a^t) & \text{if } n \neq m_{j,t}^{j}(a^t) \text{ and } \omega_{t-1} \neq j, \\ Q_{j,t}(n, a^t) - q_{j,t-1}^{j}(n, a^{t-1}) + q_{j,t-1}^{i}(n, a^{t-1}) + q_{j,t}^{j}(n, a^t) & \text{if } \omega_{t-1} = j \end{cases} \end{aligned}$$

with mod Nn_in_j. Note that we have

$$\begin{aligned} Q_{j,t}^{j}(n, a^t) &= \sum_{\substack{\tau=1: \\ n \neq m_{j,\tau}(a^\tau)}}^{t-1} \left\{ 1_{\{\omega_\tau \neq j\}} q_{j,\tau}^{j}(n, a^\tau) + 1_{\{\omega_\tau = j\}} q_{j,\tau}^{i}(n, a^\tau) \right\} + q_{j,t}^{j}(n, a^t) \\ &= \sum_{\substack{\tau=1: \\ n \neq m_{j,\tau}(a^\tau)}}^{t} \left\{ p_{j,\tau}^{i}(n, a^\tau) - q_{j,\tau-1}^{j,\omega}(n, a^{\tau-1}) \right\} \text{ by (29)} \end{aligned} \tag{30}$$

with mod Nn_in_j if $n \neq m_{j,t}^{j}(a^{t-1})$ and $\omega_t \neq j$. As will be seen below, player i is required to send $p_{j,\tau}^{i}(n, a^\tau) - q_{j,\tau-1}^{j,\omega}(n, a^{\tau-1})$ in each period τ; and in period t, player j interprets $\omega_t = j$ if and only if (30) does not hold with player i's mistake. That is, (30) corresponds to the intuitive explanation of the passcode opening the box.

Let us now define the equilibrium: The mediator sends

$$\left(\left(\begin{array}{c} r_{i,t}^{i}(a^t), r_{i,t}^{\text{punish}}(a^t), \\ m_{i,t}^{i}(a^t), \left(Q_{i,t}^{i}(n, a^t), q_{i,t}^{i,\omega}(n, a^t) \right)_{n=1}^{Nn_in_j}, \\ \left(p_{j,t}^{i}(n, a^t), q_{j,t}^{i,\omega}(n, a^t) \right)_{n=1}^{Nn_in_j} \end{array} \right)_{a^t \in A^{t-1}} \right)_{t=1}^{\infty}$$

to player i at the beginning of the game. (Recall that superscript denotes the owner of the information, so each subscript is i.)

In period 1, first, player i sends $\left(p^i_{j,1}(n,\emptyset)\right)_{n=1}^{Nn_in_j}$ to player j by cheap talk. Let $\left(\hat{p}^i_{j,1}(n,\emptyset)\right)_{n=1}^{N}$ be player i's message. Player j, who knows $m^j_{j,1}(\emptyset)$, will calculate

$$\hat{p}^i_{j,1}(\emptyset) = \hat{p}^i_{j,1}(n,\emptyset) \text{ for } n = m^j_{j,1}(\emptyset).$$

From $\hat{p}^i_j(\emptyset)$, player j calculates

$$\hat{r}_{j,1}(\emptyset) = r^j_{j,1}(\emptyset) - \hat{p}^i_{j,1}(\emptyset) \ (\bmod N n_i n_j).$$

Moreover, if there exists $n \neq m^j_{j,1}(\emptyset)$ with

$$q^j_{j,1}(n,\emptyset) \neq \hat{p}^i_{j,1}(n,\emptyset),$$

then player j will have $\hat{\omega}_1 = j$. Otherwise, $\hat{\omega}_1 \neq j$.

Second, given these variables $\hat{r}_{j,1}(\emptyset)$ and $\hat{\omega}_1$, player j takes $\hat{r}_{j,1}(\emptyset)$ if $\hat{\omega}_1 \neq j$ and $r^{\text{punish}}_{j,1}(\emptyset)$ if $\hat{\omega}_1 = j$. Finally, player j sends $\left(q^{j,\omega}_{j,1}(n,\emptyset)\right)_{n=1}^{Nn_in_j}$ to player i by cheap talk if and only if $\hat{\omega}_1 \neq j$.

Recursively, in period t, given realized a^t, player i sends the message as follows. Player j in the previous period sent $\left(q^{j,\omega}_{j,t-1}(n,a^{t-1})\right)_{n=1}^{Nn_in_j}$ if and only if $\hat{\omega}_{t-1} \neq j$. If player i receives the message, then player i calculates

$$\bar{p}^i_{j,t}(n,a^t) = p^i_{j,t}(n,a^t) - q^{j,\omega}_{j,t-1}(n,a^{t-1}). \tag{31}$$

Otherwise, player i calculates

$$\bar{p}^i_{j,t}(n,a^t) = p^i_{j,t}(n,a^t) - q^{i,\omega}_{j,t-1}(n,a^{t-1}). \tag{32}$$

Then, player i sends $\left(\bar{p}^i_{j,t}(n,a^t)\right)_{n=1}^{Nn_in_j}$ to player j by cheap talk. Let $\left(\hat{p}^i_{j,t}(n,a^t)\right)_{n=1}^{Nn_in_j}$ be player i's message. Player j, who knows $m^j_{j,t}(a^t)$, will calculate

$$\hat{p}^i_{j,t}(a^t) = \hat{p}^i_{j,t}(n,a^t) \text{ for } n = m^j_{j,t}(a^t).$$

From $\hat{p}^i_{j,t}(a^t)$, player j calculates

$$\hat{r}_{j,t}(a^t) = r^j_{j,t}(a^t) - \hat{p}^i_{j,t}(a^t) \ (\bmod N n_i n_j). \tag{33}$$

Moreover, if there exists $n \neq m^j_{j,t}(a^t)$ with

$$Q_{j,t}(n,a^t) \neq \sum_{\substack{\tau=1: \\ n \neq m_{j,\tau}(a^\tau)}}^{t} \hat{p}^i_{j,\tau}(n,a^\tau), \tag{34}$$

then player j will have $\hat{\omega}_t = j$. Otherwise, she has $\hat{\omega}_t \neq j$.

Second, given these variables $\hat{r}_{j,t}(a^t)$ and $\hat{\omega}_t$, player j takes $\hat{r}_{j,t}(a^t)$ if $\hat{\omega}_t \neq j$ and $r_{j,t}^{\text{punish}}(a^t)$ if $\hat{\omega}_t = j$. Finally, player j sends $\left(q_{j,t}^{j,\omega}(n, a^t)\right)_{n=1}^{Nn_i n_j}$ to player i by cheap talk only if $\hat{\omega}_t \neq j$.

We first check player i does not have more information about the opponent's recommendation or future recommendations. Recall that $q_{j,t}^{j}(n, a^t)$, $q_{j,t}^{j}(n, \omega_t = j, a^t)$, $q_{j,t}^{j}(n, \omega_t \neq j, a^t)$, and $q_{j,t}^{i}(n, \omega_t \neq j, a^t)$ are *i.i.d.* across players and periods. Moreover, except for (25), $\left\{q_{j,t}^{i}(n, a^t)\right\}_{n=1}^{Nn_i n_j}$ is uniformly distributed, and *i.i.d.* across players and periods. Hence, for sufficiently large N, player i cannot update the other player's information

$$\left(\left(\begin{array}{c} m_{j,t}^{j}(a^t), \left(Q_{j,t}^{j}(n, a^t), q_{j,t}^{j,\omega}(n, a^t)\right)_{n=1}^{Nn_i n_j}, \\ \left(p_{i,t}^{j}(n, a^t), q_{i,t}^{j,\omega}(n, a^t)\right)_{n=1}^{Nn_i n_j} \end{array}\right)_{a^t \in A^{t-1}}\right)_{t=1}^{\infty}.$$

Hence, player i cannot update the information about ω_t.

In addition, since player i cannot update the information about $p_{i,t}^{j}(n, a^t)$, player i does not know the future recommendation $r_{i,t}(a^t)$ by (24). Moreover, for each $\rho > 0$, for sufficiently small $p_1, p_2 > 0$, after each history, player i believes that $\omega_t = 0$ with probability no less than $1 - \rho$. Hence, knowing $r_{i,t}^{\text{punish}}(a^t)$ does not update player i's belief about the future recommendation by more than ρ. Therefore, player i does not know her own future recommendations. Moreover, since $p_{j,t}^{i}(n, a^t)$ is independent of $r_{j,t}(a^t)$, by (24), player i cannot update the information about the opponent's recommendation $r_{j,t}(a^t)$.

Let us check player i's incentive to tell the truth about $\left(\bar{p}_{j,t}^{i}(n, a^t)\right)_{n=1}^{Nn_i n_j}$. Suppose that player i has told the truth until period $t-1$. For each $\rho > 0$, for sufficiently small $p_1, p_2 > 0$, after each history, player i believes that $\omega_t = 0$ with probability no less than $1-\rho$. If player i tells the truth in period t (and follow instructions and tells the truth in the subsequent periods), then (28), (30), (31), (32), and (34) imply that $\hat{\omega}_t = 0$ if $\omega_t = 0$, and player j takes $r_{j,t}$ according to (33). On the other hand, with a high probability, $\omega_t = 0$ and so player i takes an action according to (33) with j and i reversed. By (24), therefore, players are taking actions according to $\tilde{\mu}$ with a high probability. Again, with a sufficiently small $p_1, p_2 > 0$, we have $\tilde{\mu}$ and μ sufficiently close to each other.

On the other hand, if player i tells a lie in at least one element of $\left(\bar{p}_{j,t}^{i}(n, a^t)\right)_{n=1}^{Nn_i n_j}$, then for sufficiently large N, (34) implies that $\hat{\omega}_t = j$ with a high probability. Moreover, if player i creates a situation where $\omega_t = 0$ but $\hat{\omega}_t = j$, then player j will not send $\left(q_{j,t}^{j,\omega}(n, a^t)\right)_{n=1}^{Nn_i n_j}$. Then, by (28), player i cannot know $q_{j,t}^{j,\omega}(n, a^t)$. By (30), (31), and (32), for sufficiently large N, with a high probability, player j will have (34) for subsequent periods regardless of player i's continuation strategy. That is, player j will have $\hat{\omega}_\tau = j$ and player i will be minmaxed for a sufficiently long time with a high probability. Hence, for each $\varepsilon' > \varepsilon$, for sufficiently large N, it is ε'-sequentially rational not to tell a lie in the first message.

Second, if player i deviates in actions in period t, then $\omega_{t'} = j$ for $t' > t$. To avoid being minmaxed, player i needs to create a situation

$$Q_{j,t'}(n, a^{t'}) = \sum_{\substack{\tau=1: \\ n \neq m_{j,\tau}(a^\tau)}}^{t'-1} \hat{p}^i_{j,t}(n, a^{t'}) \tag{35}$$

by (34). On the other hand, by (30), we have

$$Q^j_{j,t'}(n, a^{t'}) = \sum_{\substack{\tau=1: \\ n \neq m_{j,\tau}(a^\tau)}}^{t'-1} \left\{ 1_{\{\omega_\tau \neq j\}} q^j_{j,\tau}(n, a^\tau) + 1_{\{\omega_\tau = j\}} q^i_{j,\tau}(n, a^\tau) \right\} + q^j_{j,t'}(n, a^{t'}).$$

By (29), with $\omega_{t'} = j$, player i does not observe $q^j_{j,t'}(n, a^{t'})$ even if she knows $q^{j,\omega}_{j,t'-1}(n, a^{t'-1})$. Hence, for sufficiently large N, with a high probability, player i cannot create a situation (35). Hence, regardless of player i's continuation strategy, player j will have $\hat{\omega}_\tau = j$ for a sufficiently long time with a high probability. Hence, for each $\varepsilon' > \varepsilon$, for sufficiently large N, it is ε'-sequentially rational not to deviate.

Finally, player i is indifferent between sending $q^{j,\omega}_{j,t}(n, a^t)$ correctly or not, since it does not affect player j's actions. Therefore, this is an ε'-sequential equilibrium.