

# Efficient Outcomes in Repeated Games with Limited Monitoring

Mihaela van der Schaar · Yuanzhang  
Xiao · William Zame

Received: September 1, 2014 / Accepted: June 1, 2015

**Abstract** The Folk Theorem for infinitely repeated games with imperfect public monitoring implies that for a general class of games, nearly efficient payoffs can be supported in perfect public equilibrium (PPE) provided the monitoring structure is sufficiently rich and players are arbitrarily patient. This paper shows that for stage games in which actions of players interfere strongly with each other, exactly efficient payoffs can be supported in PPE even when the monitoring structure is not rich and players are not arbitrarily patient. The class of stage games we study abstracts many environments including resource sharing.

**Keywords** repeated games · imperfect public monitoring · perfect public equilibrium · efficient outcomes · repeated resource allocation · repeated partnership · repeated contest

**Mathematics Subject Classification (2000)** JEL C72 · C73 · D02

---

This research was supported by National Science Foundation (NSF) Grants No. 0830556, (van der Schaar, Xiao) and 0617027 (Zame) and by the Einaudi Institute for Economics and Finance (Zame). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agency.

---

Mihaela van der Schaar  
Electrical Engineering Department, UCLA  
E-mail: mihaela@ee.ucla.edu

Yuanzhang Xiao  
Electrical Engineering Department, UCLA  
E-mail: xyz.xiao@gmail.com

William Zame  
Department of Economics, UCLA  
Tel.: (+1) 310-985-3091  
E-mail: zame@econ.ucla.edu

## 1 Introduction

The Folk Theorem for infinitely repeated games with imperfect public monitoring (Fudenberg, Levine, and Maskin (1994); henceforth FLM) implies that, under technical (full-rank) conditions on the stage game, nearly efficient payoffs can be supported in perfect public equilibrium (PPE) under the assumptions that players are arbitrarily patient (i.e., the common discount factor tends to 1) and the monitoring structure is sufficiently rich. For general stage games, the restriction to nearly efficient payoffs and the assumptions that players are arbitrarily patient and that the monitoring structure is sufficiently rich are all necessary: It is easy to construct stage games and imperfect monitoring structures for which exactly efficient payoffs cannot be supported for any discount factor (less than 1) and nearly efficient payoffs cannot be supported for discount factors bounded away from 1. And Radner, Myerson, and Maskin (1986) construct a repeated partnership scenario in which players see only two signals – success or failure – and even nearly efficient payoffs cannot be supported even when players are arbitrarily patient.

This paper shows that, for a large and important class of stage games, *exactly* efficient payoffs are supportable in PPE even when the monitoring structure is very limited. The stage games we consider are those that arise in many common and important settings in which the actions of players *interfere* with each other. The paradigmatic setting is that of sharing a resource that can be efficiently accessed by only one player at a time – so that efficient sharing requires alternation over time – but as we illustrate by examples, the same interference phenomenon may be present in partnership games and in contests (and surely in many other scenarios). We focus on monitoring structures with only two signals, this is a restriction but it is stark and easy to understand, and also very natural: in the partnership scenario of Radner, Myerson, and Maskin (1986) for instance, the signal is the success or failure of the partnership interaction. (As we discuss later, an additional reason for focusing on simple monitoring structures is that the signal does not necessarily arise directly from the actions of the players as in or from the interactions of the players with the market as in Green and Porter (1984) but rather must be provided by some outside agency, which may face constraints and costs on what it can observe and what it can communicate to the players.) A feature of our work that we think is important in any realistic setting is that it is *constructive*: given an efficient target payoff profile, we explicitly identify the degree of patience players must exhibit in order that the target payoff be achievable in PPE and we provide a simple explicit algorithm that allows each player to compute (based on public information) its equilibrium action in each period. For games with two players, we show that the set of efficient payoffs that can be supported as a PPE is independent of the discount factor provided the discount factor is above some threshold.<sup>1</sup>

---

<sup>1</sup> Mailath, Obara, and Sekiguchi (2002) establish a similar result for the repeated Prisoner's Dilemma with perfect monitoring; Athey and Bagwell (2001) establish a similar re-

We abstract what we see as the essential common features of a variety of scenarios by two assumptions about the stage game. The first is that for each player  $i$  there is a unique action profile  $\tilde{\mathbf{a}}^i$  that  $i$  most prefers. (In the resource sharing scenario,  $\tilde{\mathbf{a}}^i$  would typically be the profile in which only player  $i$  accesses the resource; in the partnership scenario it would typically be the profile in which player  $i$  free rides.) The second is that for every action profile  $\mathbf{a}$  that is *not* in the set  $\{\tilde{\mathbf{a}}^i\}$  of preferred action profiles the corresponding utility profile  $u(\mathbf{a})$  lies *below* the hyperplane  $H$  spanned by the utility profiles  $\{u(\tilde{\mathbf{a}}^i)\}$ . (This corresponds to the idea that player actions interfere with each other, rather than complementing each other.) As usual in this literature, we assume that players do not observe the profile  $\mathbf{a}$  of actions but rather only some signal  $y \in Y$  whose distribution  $\rho(y|\mathbf{a})$  depends on the true profile  $\mathbf{a}$ . We depart from the literature by assuming that the set  $Y$  consists of only two signals and that (profitable) single-player deviations from any of the preferred action profiles  $\tilde{\mathbf{a}}^i$  can be statistically distinguished from conformity with  $\tilde{\mathbf{a}}^i$  by altering the probability distribution on signals *in the same direction*. (But we do not assume that different deviations from  $\tilde{\mathbf{a}}^i$  can be distinguished *from each other*.) For further comments, see the examples in Section 3.

To help understand the commonplace nature of our problem and assumptions, we offer three examples: a repeated partnership game, a repeated contest, and a repeated resource sharing game. In the repeated partnership game, the signal arises from the state of the market. In the repeated contest, signals arise from direct observation of the outcome of the contest or from information provided by the agency that conducts the contest. In this setting there is a natural choice of signal structures and hence of the amount of information to provide, and this choice affects the possibility of efficient PPE. In the repeated resource sharing game, signals are provided by an outside agency. In this setting there is again a natural choice of signal structures and the choice affects the distribution of information provided but not the amount, and so has quite a different effect on the possibility of efficient PPE. As we discuss, the agency's choice between signal structures will most naturally be determined by the agency's objectives; simulations show that different objectives are best served by different signal structures.

Our constructions build on the framework of Abreu, Pearce, and Stacchetti (1990) (hereafter APS) and in particular on the machinery of *self-generating sets*. APS show that every payoff in a self-generating set can be supported in a perfect public equilibrium, so it is no surprise that we prove our main result (Theorem 1) by constructing appropriate self-generating sets of a particularly simple form. A technical result that seems of substantial interest in itself (Theorem 2) provides necessary and sufficient conditions that sets of this form be self-generating. Our construction provides an explicit algorithm for computing PPE strategies using continuation values in the constructed self-generating sets. Because all continuation payoffs lie in the specified self-

---

sult for equilibrium payoffs of two-player symmetric repeated Bertrand games. Mailath and Samuelson (2006) present examples with restricted signal structures.

generating set, the equilibria we construct have the property that each player is guaranteed at least a specific security level following every history. Because all continuation payoffs are efficient, the equilibria we construct are renegotiation-proof following every history: players would never unanimously agree to follow an alternative strategy profile. (Fudenberg, Levine, and Takahashi (2007) – henceforth FLT – emphasize the same point.)

The literature on repeated games with imperfect public monitoring is quite large – much too large to survey here; we refer instead to Mailath and Samuelson (2006) and the references therein. However, explicit comparisons with two papers in this literature may be especially helpful. As we have noted, FLM consider general stage games (subject to some technical conditions) but assume that the monitoring structure is rich – in particular that there are many signals – and only establish the existence of nearly-efficient PPE. Moreover, FLM require discount factors arbitrarily close to 1 in order to obtain PPE that are arbitrarily close to efficient. By contrast, we restrict to a (natural and important) class of stage games, we require only two signals *even if action spaces are infinite*, and we obtain exactly efficient PPE. FLT is much closer to the present work. FLT fix Pareto weights  $\lambda_1, \dots, \lambda_n$  for which the feasible set  $X$  lies weakly below the hyperplane  $H = \{x \in \mathbb{R}^n : \sum \lambda_i x_i = A\}$ , so that the intersection  $V = H \cap X$  consists of exactly efficient payoff vectors. As do we, FLT ask what vectors in  $V$  can be achieved as a PPE of the infinitely repeated game. They identify the largest (compact convex) set  $Q \subset V$  with the property that every target vector  $v \in \text{int}Q$  (the relative interior of  $Q$  with respect to  $H$ ) can be achieved in a PPE of the infinitely repeated game for *some* discount factor  $\delta(v) < 1$ . However, for general stage games and general monitoring structures, the set  $Q$  may be empty; FLT do not offer conditions that guarantee that  $Q$  is not empty. Moreover (as do FLM), FLT focus on what can be achieved when players are arbitrarily patient; even when  $Q$  is not empty, they do not identify any PPE for any *given* discount factor  $\delta < 1$ . We give specific conditions that guarantee that  $Q$  is not empty and provide explicit and computable PPE strategies for given discount factors. For games with two players, FLT find a sufficient condition that there be no efficient PPE for any discount factor; we find a (sharper) sufficient *and* necessary condition and we show that the set of efficient payoffs that *can* be supported as a PPE is independent of the discount factor provided the discount factor is above some threshold. See Section 6 for additional comparisons with results in the unpublished working paper version of FLT.

At the risk of repetition, we want to emphasize several features of our results. The first is that we do not assume discount factors are arbitrarily close to 1; rather we give explicit sufficient conditions on the discount factors (and on the other aspects of the environment) to guarantee the existence of PPE. The importance of this seems obvious in all environments – especially since the discount factor encodes both the innate patience of players *and* the probability that the interaction continues. The second is that we assume only two signals, even when action spaces are infinite. Again, the importance of this seems obvious in all environments, but especially in those in which signals are

not generated by some exogenous process but must be provided. (In the latter case it seems obvious – and in practice may be of supreme importance – that the agency providing signals may wish or need to choose a simple information structure that employs a small number of signals, saving on the cost of observing the outcome of play and on the cost of communicating to the agents. More generally, there may be a trade-off between the efficiency obtainable with a finer information structure and the cost of using that information structure.) Finally, we provide a simple distributed algorithm that enables each player to calculate its equilibrium play online, in real-time, period by period (not necessarily at the beginning of the game).

Following this Introduction, Section 2 presents the formal model; Section 3 presents three examples that illustrate the model. Section 4 presents the main theorem (Theorem 1) on supportability of efficient outcomes in PPE. Section 5 presents the more technical result (Theorem 2) characterizing efficient self-generating sets. Section 6 specializes to the case of two players (Theorem 3). Section 7 concludes. We relegate all proofs to the Appendix.

## 2 Model

We first describe the general structure of repeated games with imperfect public monitoring; our description is parallel to that of FLM and Mailath and Samuelson (2006) (henceforth MS). Following the description we formulate the assumptions for the specific class of games we treat.

### 2.1 Stage Game

The *stage game*  $G$  is specified by :

- a set  $N = \{1, \dots, n\}$  of *players*
- for each player  $i$ 
  - a compact metric space  $A_i$  of *actions*
  - a continuous *utility function*  $u_i : \mathbf{A} = A_1 \times \dots \times A_n \rightarrow \mathbb{R}$

### 2.2 Public Monitoring Structure

The public monitoring structure is specified by

- a finite set  $Y$  of public *signals*
- a continuous mapping  $\rho : \mathbf{A} \rightarrow \Delta(Y)$

As usual, we write  $\rho(y|\mathbf{a})$  for the probability that the public signal  $y$  is observed when players choose the action profile  $\mathbf{a} \in \mathbf{A}$ .

### 2.3 The Repeated Game with Imperfect Public Monitoring

In the repeated game, the stage game  $G$  is played in every period  $t = 0, 1, 2, \dots$ . If  $Y$  is the set of public signals, then a *public history* of length  $t$  is a sequence  $(y^0, y^1, \dots, y^{t-1}) \in Y^t$ . We write  $\mathcal{H}(t)$  for the set of public histories of length  $t$ ,  $\mathcal{H}^T = \bigcup_{t=0}^T \mathcal{H}(t)$  for the set of public histories of length at most  $T$  and  $\mathcal{H} = \bigcup_{t=0}^{\infty} \mathcal{H}(t)$  for the set of all public histories of all finite lengths. A *private history* for player  $i$  includes the public history and the actions taken by player  $i$ , so a *private history* of length  $t$  is a sequence  $(a_i^0, y^0; \dots, a_i^{t-1}, y^{t-1}) \in \mathbf{A}_i^t \times Y^t$ . We write  $\mathcal{H}_i(t)$  for the set of  $i$ 's private histories of length  $t$ ,  $\mathcal{H}_i^T = \bigcup_{t=0}^T \mathcal{H}_i(t)$  for the set of  $i$ 's private histories of length at most  $T$  and  $\mathcal{H}_i = \bigcup_{t=0}^{\infty} \mathcal{H}_i(t)$  for the set of  $i$ 's private histories of all finite lengths.

A *pure strategy* for player  $i$  is a mapping from all private histories into player  $i$ 's set of actions  $\sigma_i : \mathcal{H}_i \rightarrow A_i$ . A *public strategy* for player  $i$  is a pure strategy that is independent of  $i$ 's own action history; equivalently, a mapping from public histories to  $i$ 's pure actions  $\sigma_i : \mathcal{H} \rightarrow A_i$ .

We assume as usual that all players discount future utilities using the same discount factor  $\delta \in (0, 1)$  and we use long-run averages: if  $\{u^t\}$  is the stream of expected utilities then the vector of long-run average utilities is  $(1 - \delta) \sum_{t=0}^{\infty} \delta^t u^t$ . (Note that we do not discount date 0 utilities). A strategy profile  $\sigma : \mathcal{H}_1 \times \dots \times \mathcal{H}_n \rightarrow \mathbf{A}$  induces a probability distribution over public and private histories and hence over *ex ante* utilities. We abuse notation and write  $u(\sigma)$  for the vector of expected (with respect to this distribution) long-run average *ex ante* utilities when players follow the strategy profile  $\sigma$ .

As usual a strategy profile  $\sigma$  is an *equilibrium* if each player's strategy is optimal given the strategies of others. A strategy profile is a *public equilibrium* if it is an equilibrium and each player uses a public strategy; it is a *perfect public equilibrium (PPE)* if it is a public equilibrium following every public history. Note that if the signal distribution  $\rho(y|\mathbf{a})$  has full support for every action profile  $\mathbf{a}$  then every public history always occurs with strictly positive probability so perfect public equilibrium coincides with public equilibrium. Keeping the stage game  $G$  and the monitoring structure  $Y, \rho$  fixed, we write  $\mathcal{E}(\delta)$  for the set of long-run average payoffs that can be achieved in a PPE of the infinitely repeated game when the discount factor is  $\delta < 1$ .

### 2.4 Interpretation

We interpret payoffs in the stage game as *ex ante* payoffs. Note that this interpretation allows for the possibility that each player's *ex post*/realized payoff depends on the actions of all players and the realization of the public signal – and perhaps on the realization of some other random event (see the examples).

Of course players do not observe *ex ante* payoffs – they observe only their own actions and the public signal.<sup>2</sup>

In our formulation, which restricts players to use public strategies, we tacitly assume that players make no use of any information other than that provided by the public signal; in particular, players make no use of information that might be provided by the realized utility they experience each period. As discussed in FLM and MS, this assumption admits a number of possible interpretations; one is that players do not observe their realized period utilities, but only the total realized utility at the termination of the interaction.

It is important to keep in mind that if players other than player  $i$  use a public strategy, then it is always a best response for player  $i$  to use a public strategy (MS, Lemma 7.1.1). Moreover, requiring agents to use public strategies in equilibrium but allowing arbitrary deviation strategies (as we do) means that *fewer* outcomes can be supported in equilibrium than if we allowed agents to use arbitrary strategies in equilibrium. Since our intent is to show that efficient outcomes can be supported, restricting to perfect public equilibrium makes our task *more difficult*.

## 2.5 Games with Interference

To this point we have described a very general setting; we now impose additional assumptions – first on the stage game and then on the information structure – that we exploit in our results.

Set  $U = \{u(\mathbf{a}) \in \mathbb{R}^n : \mathbf{a} \in \mathbf{A}\}$  and let  $X = \text{co}(U)$  be the closed convex hull of  $U$ . For each  $i$  set

$$\begin{aligned}\tilde{v}_i^i &= \max_{\mathbf{a} \in \mathbf{A}} u_i(\mathbf{a}) \\ \tilde{\mathbf{a}}^i &= \arg \max_{\mathbf{a} \in \mathbf{A}} u_i(\mathbf{a})\end{aligned}$$

Compactness of the action space  $\mathbf{A}$  and continuity of utility functions  $u_i$  guarantee that  $U$  and  $X$  are compact, that  $\tilde{v}_i^i$  is well-defined and that the arg max is not empty. For convenience, we assume that the arg max is a singleton; i.e., the maximum utility  $\tilde{v}_i^i$  for player  $i$  is attained at a *unique* strategy profile  $\tilde{\mathbf{a}}^i$ .<sup>3</sup> We refer to  $\tilde{\mathbf{a}}^i$  as  $i$ 's *preferred action profile* and to  $\tilde{v}^i = u(\tilde{\mathbf{a}}^i)$  as  $i$ 's *preferred utility profile*. In the context of resource sharing,  $\tilde{\mathbf{a}}^i$  will be the (unique) action profile at which agent  $i$  has optimal access to the resource and other players have none; in some other contexts,  $\tilde{\mathbf{a}}^i$  will be the (unique) action profile at which  $i$  exerts effort and other players exert none. For this reason, we will often say that  $i$  is *active* at the profile  $\tilde{\mathbf{a}}^i$  and other players are *inactive*. (However we caution the reader that in the repeated partnership game of Example

<sup>2</sup> Although it is often assumed that each player's *ex post*/realized payoff depends only on the its own action and the public signal, FLM explicitly allow for the more general setting we consider here.

<sup>3</sup> The assumption of uniqueness could be avoided, at the expense of some technical complication.

1,  $\tilde{\mathbf{a}}^i$  is the action profile at which player  $i$  is free riding and his partner is exerting effort.) Set  $\tilde{\mathbf{A}} = \{\tilde{\mathbf{a}}^i\}$  and  $\tilde{V} = \{\tilde{v}^i\}$  and write  $V = \text{co}(\tilde{V})$  for the convex hull of  $\tilde{V}$ . Note that  $X$  is the convex hull of the set of vectors that can be achieved – for *some* discount factor – as long-run average *ex ante* utilities of repeated plays of the game  $G$  (not necessarily equilibrium plays of course) and that  $V$  is the convex hull of the set of vectors that can be achieved – for *some* discount factor – as long-run average *ex ante* utilities of repeated plays of the game  $G$  in which only actions in  $\tilde{\mathbf{A}}$  are used. We refer to  $X$  as the set of *feasible payoffs* and to  $V$  as the set of *efficient payoffs*.<sup>4</sup>

We abstract the motivating class of resource allocation problems by imposing a condition on the set of preferred utility profiles.

**Assumption 1** *The set  $\{\tilde{v}^i\}$  of preferred utility vectors is a linearly independent set and there are (Pareto) weights  $\lambda_1, \dots, \lambda_n > 0$  such that  $\sum_j \lambda_j \tilde{v}_j^i = 1$  for each  $i$  and  $\sum_j \lambda_j u_j(\mathbf{a}) < 1$  for each  $\mathbf{a} \in \mathbf{A}, \mathbf{a} \notin \tilde{\mathbf{A}}$ . (Thus the set  $H = \{x \in \mathbb{R}^n : \sum_j \lambda_j x_j = 1\}$  is a hyperplane, payoffs in  $\tilde{V}$  lie in  $H$  and all pure strategy payoffs not in  $\tilde{V}$  lie strictly below  $H$ . That the sum  $\sum_j \lambda_j \tilde{v}_j^i$  is 1 is just a normalization.)*

## 2.6 Assumptions on the Monitoring Structure

As noted in the Introduction, we assume there are only two signals and that profitable deviations from the profiles  $\tilde{\mathbf{a}}^i$  exist and are statistically detectable in a particularly simple way.

**Assumption 2** *The set  $Y$  contains precisely two signals  $g, b$  (good, bad).*

**Assumption 3** *For each  $i \in N$  and each  $j \neq i$  there is an action  $a_j \in A_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$ . Moreover,*

$$a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i) \Rightarrow \rho(g|a_j, \tilde{\mathbf{a}}_{-j}^i) < \rho(g|\tilde{\mathbf{a}}_j^i, \tilde{\mathbf{a}}_{-j}^i)$$

That is, given that other players are following  $\tilde{\mathbf{a}}^i$ , any strictly profitable deviation by player  $j$  strictly reduces the probability that the good signal  $g$  is observed and so strictly increases the probability that the bad signal  $b$  is observed.<sup>5,6</sup>

<sup>4</sup> This is a slight abuse of terminology. Assumption 1 below is that  $V$  is the intersection of the set of feasible payoffs with a bounding hyperplane, so every payoff vector in  $V$  is Pareto efficient and yields maximal weighted social welfare and other feasible payoffs yield lower weighted social welfare – but other feasible payoffs might also be Pareto efficient.

<sup>5</sup> The assumption that the same signals are good/bad independently of the identity of the active player  $i$  is made only to reduce the notational burden. The interested reader will easily check that all our arguments allow for the possibility that which signal is good and which is bad depend on the identity of the active player.

<sup>6</sup> The restriction to two signals is not entirely innocuous. If there were more than two signals, the conditions identified in Theorem 2 will continue to be sufficient for a set to be self-generating but may no longer be necessary. Moreover, exploiting a richer set of signals may lead to a larger set of PPE; see the discussions following Examples 2 and 3.



**Table 1** Partnership Game – Realized Payoffs

	$g$	$b$
$E$	$g/2 - e$	$b/2 - e$
$S$	$g/2$	$b/2$

Assumption 3 guarantees that all profitable single player deviations from  $\tilde{\mathbf{a}}^i$  alter the signal distribution in the *same direction* – although perhaps not to the same extent. We allow for the possibility that non-profitable deviations may not be detectable in the same way – perhaps not detectable at all.

### 3 Examples

The assumptions we have made – about the structure of the game and about the information structure – are far from innocuous, but they apply in a wide variety of interesting environments. Here we describe three simple examples which motivate and illustrate the assumptions we have made and the conclusions to follow.

The first example is a repeated partnership, very much in the spirit of an example in MS (Section 7.2) but with a twist.

#### Example 1: Repeated Partnership

Each of two partners can choose to exert costly effort  $E$  or shirk  $S$ . Realized output can be either Good  $g$  or Bad  $b$  ( $g > b > 0$ ), and depends stochastically on the effort of the partners. Realized individual payoffs as a function of actions and realized output are shown in Table 1.

In contrast to MS, we assume that if both players exert effort they interfere with each other. Output follows the distribution

$$\rho(g|\mathbf{a}) = \begin{cases} p & \text{if } \mathbf{a} = (E, S) \text{ or } (S, E) \\ q & \text{if } \mathbf{a} = (E, E) \\ r & \text{if } \mathbf{a} = (S, S) \end{cases}$$

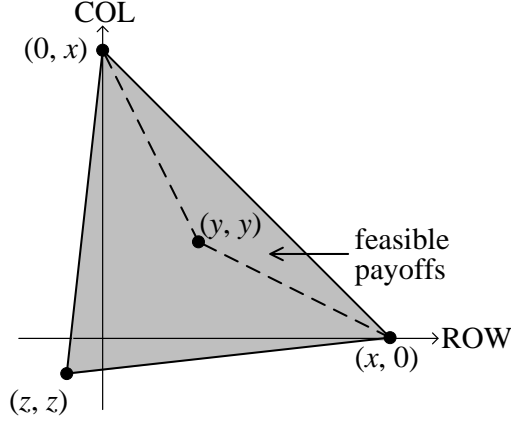
where  $p, q, r \in (0, 1)$  and  $p > q > r$ . The signal is most likely to be  $g$  (high output) if *exactly* one partner exerts effort.

The *ex ante* payoffs can be calculated from the data above; it is convenient to normalize so that the *ex ante* payoff to the player who exerts effort when his partner shirks is 0:  $(1/2)[pg + (1 - p)b] - e = 0$ . With this normalization, the *ex ante* game matrix  $G$  is shown in Table 2; we assume parameters are such that  $x > 2y > 0 > z$  (we leave it to the reader to calculate the values of  $x, y, z$  in terms of output  $g, b$  and probabilities  $p, q, r$ ).

It is easily checked that the stage game and monitoring structure satisfy our assumptions:  $(S, E)$  is the preferred profile for ROW and  $(E, S)$  is the

**Table 2** Partnership Game – Ex Ante Payoffs

	$E$	$S$
$E$	$(z, z)$	$(0, x)$
$S$	$(x, 0)$	$(y, y)$

**Fig. 1** Feasible Region for the Repeated Partnership Game

preferred profile for COL. Figure 1 shows the feasible region for the repeated partnership game.<sup>7</sup>

As we will show in Section 6, we can completely characterize the most efficient outcomes that can be achieved in a PPE. For  $x \leq 2p/(p-r)y$ , there is no efficient PPE payoff for *any* discount factor  $\delta \in (0, 1)$ . For  $x > 2p/(p-r)y$ , set

$$\delta^* = \frac{1}{1 + \left( \frac{x - \frac{p}{p-r} \cdot 2y}{x + \frac{1-p}{p-r} \cdot 2y} \right)}$$

It follows from Theorem 3 that if  $\delta \geq \delta^*$  then

$$\mathcal{E}(\delta) = \{(v_1, v_2) : v_1 + v_2 = x; v_i \geq p/(p-r)y\}$$

Note that the set of efficient PPE outcomes *does not* increase as  $\delta \rightarrow 1$ ; patience is rewarded – but only up to a point.

If we identify the monitoring technology with the probabilities  $p, q, r$  we should note that different monitoring technologies provide different information, but that there may not be any natural ordering in the sense of Blackwell informativeness (for instance, if we are given alternative probabilities  $p', q', r'$  with  $|p - .5| < |p' - .5|$  but  $|r - .5| > |r' - .5|$  then the monitoring technologies are not comparable in the sense of Blackwell informativeness), so that the results of Kandori Kandori (1992) do not apply.

<sup>7</sup> Note that if  $x < 2y$  then the the stage game fails Assumption 1; in particular, some payoffs in the convex hull of the preferred profiles  $(E, S), (S, E)$  are not Pareto optimal.

**Example 2: Repeated Contest**

In each period, a set of  $n \geq 2$  players competes for a single indivisible prize that each of them values at  $R > 0$ . Winning the competition depends (stochastically) on the effort exerted by each player. Each agent's effort interferes with the effort of others and there is always some probability that no one wins (the prize is not awarded) independently of the choice of effort levels. The set of  $i$ 's effort levels/actions is  $A_i = [0, 1]$ . If  $\mathbf{a} = (a_i)$  is the vector of effort levels then the probability agent  $i$  wins the competition and obtains the prize is

$$\text{Prob}(i \text{ wins} | \mathbf{a}) = a_i \left( \eta - \kappa \sum_{j \neq i} a_j \right)^+$$

where  $\eta, \kappa \in (0, 1)$  are parameters, and  $(\cdot)^+ \triangleq \max\{\cdot, 0\}$ . That  $\eta < 1$  reflects that there is always some probability the prize is not awarded;  $\kappa$  measures the strength of the interference. Notice that competition is destructive: if more than one agent exerts effort that lowers the probability that *anyone* wins. Utility is separable in reward and effort; effort is costly with constant marginal cost  $c > 0$ . To avoid trivialities and conform with our Assumptions we assume  $R\eta > c$ ,  $(\eta + \kappa)^2 < 4\kappa$ , and  $\kappa > \frac{\eta}{2}$ .

We assume that, at the end of each period of play, players observe (or are told) only whether or not the prize was awarded (but not to whom). So the signal space is  $Y = \{g, b\}$ , where  $g$  is interpreted as the announcement that the prize was awarded and  $b$  is interpreted as the announcement that the prize was not awarded.<sup>8</sup>

The *ex ante* expected utilities for the stage game  $G$  are given by

$$u_i(\mathbf{a}) = a_i \left( \eta - \kappa \sum_{j \neq i} a_j \right)^+ R - ca_i$$

The signal distribution is defined by

$$\rho(g | \mathbf{a}) = \sum_i a_i \left( \eta - \kappa \sum_{j \neq i} a_j \right)^+$$

Straightforward calculations show our assumptions are satisfied. Player  $i$ 's preferred action profile  $\tilde{\mathbf{a}}^i$  has  $\tilde{a}_i^i = 1$  and  $\tilde{a}_j^i = 0$  for  $j \neq i$ :  $i$  exerts maximum effort, others exert none. Note that this does not guarantee that  $i$  wins the prize – the prize may not be awarded – but the effort profiles  $\tilde{\mathbf{a}}^i$  are precisely those that maximize the probability that *someone* wins the prize.

We have assumed that, in each period, players learn whether or not someone wins the competition but *do not learn the identity of the winner*. We might

<sup>8</sup> Note that realized payoffs depend on who actually wins the prize, not only on the profile of actions and the announcement.

consider an alternative monitoring structure in which the players *do learn the identity of the winner*. To see why this matters, suppose that a strategy profile  $\sigma$  calls for  $\tilde{\mathbf{a}}^i$  to be played after a particular history. If all players follow  $\sigma$  then only player  $i$  exerts non-zero effort so only two outcomes can occur: either player  $i$  wins or no one wins. If player  $j \neq i$  deviates by exerting non-zero effort, a third outcome can occur:  $j$  wins. With either monitoring structure, it is possible for the players to detect (statistically) that *someone* has deviated – the probability that *someone* wins goes down – but with the second monitoring structure it is also possible for the players to detect (statistically) *who* has deviated – because the probability that the deviator wins becomes positive. Hence, with the first monitoring structure all deviations must be “punished” in the same way, but with the second monitoring structure, “punishments” can be tailored to the deviator. If punishments can be “tailored” to the deviator then punishments can be more severe; if punishments can be more severe it may be possible to sustain a wider range of PPE. In short: the monitoring structure matters.

But the monitoring structure is not arbitrary: players will not learn the identity of the winner unless they can observe it directly – which might or might not be possible in a given scenario – or they are informed of it by an outside agency – which requires the outside agency to reveal additional information. This is information the agency conducting the contest would possess – but whether or not this is the information the agency would wish – or be permitted – to reveal would seem to depend on the environment. A similar point is made more sharply in the final example below.

### Example 3: Repeated Resource Sharing

We consider a very common communication scenario.  $N \geq 3$  users (players) send information packets through a common server. The server has a nominal capacity of  $\chi > 0$  (packets per unit time) but the capacity is subject to random shocks so the actually realized capacity in a given period is  $\chi - \varepsilon$ , where the random shock  $\varepsilon$  is distributed in some interval  $[0, \bar{\varepsilon}]$  with (known) uniform distribution  $\nu$ . In each period, each player chooses a packet rate (packets per unit time)  $a_i \in A_i = [0, \chi]$ . This is a well-studied problem; assuming that the players’ packets arrive according to a Poisson process, the whole system can be viewed as what is known as an *M/M/1 queue*; see Bharath-Kumar and Jaffe (1981) for instance. It follows from the standard analysis that if  $\varepsilon$  is the realization of the shock then packet deliveries will be subject to a *delay* of

$$D(\mathbf{a}, \varepsilon) = \begin{cases} 1/(\chi - \varepsilon - \sum_{i=1}^N a_i) & \text{if } \sum_{i=1}^N a_i < \chi - \varepsilon \\ \infty & \text{if } \sum_{i=1}^N a_i \geq \chi - \varepsilon \end{cases}$$

Given the delay  $D$ , each player’s realized utility is its “power” – the ratio of the  $p$ -th power of its own packet rate to the delay:

$$u_i^*(\mathbf{a}, D) = a_i^p / D$$

The exponent  $p > 0$  is a parameter that represents trade-off between rate and delay.<sup>9</sup> (If delay is infinite utility is 0.)

The server is monitored by an agency that does not observe packet rates but can measure the delay; however, measurement is costly and subject to error. We assume the agency reports to the players, not the measured delay, but whether it is above or below a chosen threshold  $D_0$ . Thus  $Y = \{g, b\}$  where  $g$  is interpreted as “delay was low (below  $D_0$ )” and  $b$  is interpreted as “delay was high (above  $D_0$ ).”

Each player  $i$ 's ex-ante payoff is

$$u_i(\mathbf{a}) = \begin{cases} a_i^p (\chi - \frac{\bar{\varepsilon}}{2} - \sum a_j) & \text{if } \sum a_j \leq \chi - \bar{\varepsilon} \\ a_i^p (\chi - \sum a_j) \frac{\chi - \sum a_j}{2\bar{\varepsilon}} & \text{if } \chi - \bar{\varepsilon} < \sum a_j < \chi \\ 0 & \text{otherwise} \end{cases}$$

and the distribution of signals is

$$\rho(g|\mathbf{a}) = \int_0^{\chi - \sum a_j - \frac{1}{D_0}} d\nu(x) = \frac{[\chi - \sum a_j - \frac{1}{D_0}]_0^{\bar{\varepsilon}}}{\bar{\varepsilon}},$$

where  $[x]_a^b \triangleq \min\{\max\{x, a\}, b\}$  is the projection of  $x$  in the interval  $[a, b]$  and all summations are taken over the range  $j = 1, \dots, N$ . As noted,  $g$  is the “good” signal: deviation from any preferred action profile increases the probability of realized delay, hence increases the probability of measured delay, and reduces the probability that reported delay will be below the chosen threshold.

Because players do not observe delay directly, the signal of delay must be provided. It is natural to suppose this signal is provided by some agency, which must choose the technology by which it observes delay and the threshold  $D_0$  “low delay” and “high delay”. These choices will presumably be made according to some objective – but different objectives will lead to different choices of  $D_0$  and there is no “obviously correct” objective.<sup>10</sup> (It is important to note that a higher/lower threshold  $D_0$  *does not* correspond to more/less information, so the choice of  $D_0$  *is not* the choice of how much information to reveal.)

This can be seen clearly in numerical results for a special case. Set capacity  $\chi = 1$  and  $\bar{\varepsilon} = 0.3$ . We consider two possible objectives.

- The agency chooses the threshold  $D_0$  to minimize the discount factor  $\delta$  for which *some* efficient sharing can be supported in a PPE.
- The agency chooses the threshold  $D_0$  to maximize the *set of efficient payoffs* that can be supported in PPE for *some* discount factor  $\delta$ . This is a somewhat imprecise objective; to make it precise, set

$$V(\eta) = \{v \in V : v_i \geq \eta \bar{v} \text{ for each } i\}$$

<sup>9</sup> In order to guarantee our assumptions are satisfied we assume  $\bar{\varepsilon} \leq \frac{2}{2+p}\chi$ .

<sup>10</sup> Presumably the agency would prefer a more accurate measurement technology – but such a technology would typically be more costly to employ.

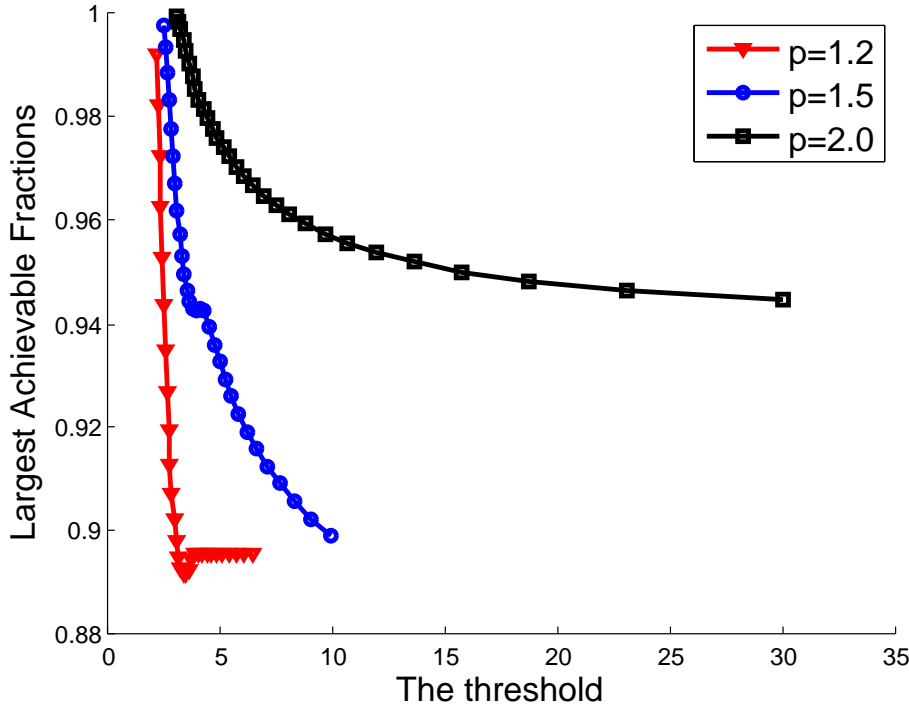


Fig. 2 Largest Achievable Fraction  $1 - \eta$  as a Function of Threshold  $D_0$ .

where  $\tilde{v}$  is the utility of each player's most preferred action and  $\eta \in [0, 1]$ . Note that  $V(\eta) \subset V(\eta')$  if  $\eta' < \eta$  so to maximize the set of efficient payoffs that can be supported in PPE for *some* discount factor  $\delta$ , the agency should choose  $D_0$  so that  $V(\eta) \subset \mathcal{E}(\delta)$  for some  $\delta$  and the smallest possible  $\eta$ .

Figures 2 and 3 (which are generated from simulations) display the relationship between the threshold  $D_0$ , the smallest  $\delta$  and the smallest  $\eta$  for several values of the exponent  $p$ . The tension between the criteria for choosing the threshold  $D_0$  can be seen most clearly when  $p = 1.2$ : to make it easiest to achieve *many* efficient outcomes the agency should choose a small threshold, but to make it easiest to achieve *some* efficient outcome the agency should choose a large threshold.

We noted in Example 1 that different monitoring structures provide different information, but that there may not be any natural ordering in the sense of Blackwell informativeness, so that the results of Kandori (1992) do not apply. In the current Example, note that a different choices of threshold  $D_0$  yield different information, but that higher thresholds are neither more nor less informative in the sense of Blackwell.

A final remark about this Example may be useful. We have assumed throughout that players do not condition on their realized utility but it is worth noting that in this case, even if players did condition on their real-

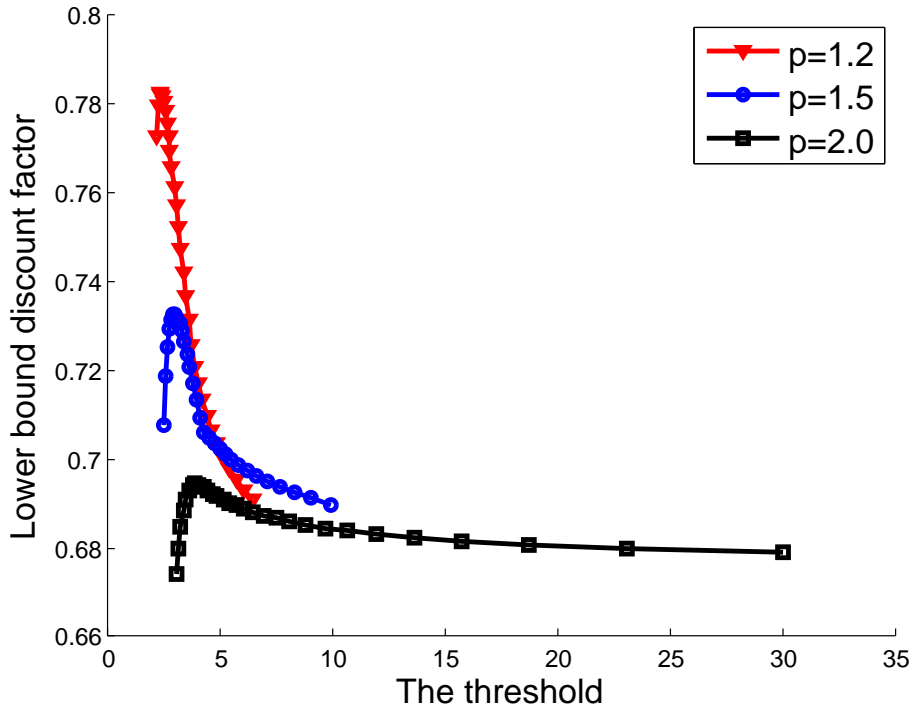


Fig. 3 Smallest Achievable Discount Factor  $\delta$  as a Function of Threshold  $D_0$

ized utility monitoring would still be imperfect. While players who transmit (choose packet rates greater than 0) could back out realized delay, players who do not transmit cannot back out realized delay and must therefore rely on the announcement of delay to know how to behave in the next period. Hence these announcements serve to keep players on the same informational page.

#### 4 Perfect Public Equilibria

From this point on, we consider a fixed stage game  $G$  and monitoring structure  $Y, \rho$  and maintain the notation and assumptions of Section 2. For fixed  $\delta \in (0, 1)$  we write  $\mathcal{E}(\delta)$  for the set of (long-run average) payoffs that can be achieved in a PPE when the discount factor is  $\delta$ . Our goal is to find conditions – on payoffs, signal probabilities and discount factor – that enable us to construct PPE that achieve efficient payoffs with some degree of sharing among all players. In other words, we are interested in conditions that guarantee that  $\mathcal{E}(\delta) \cap \text{int } V \neq \emptyset$ .

In order to write down the conditions we need, we first introduce some notions and notations. The first notions are two measures of the profitability of deviations; these play a prominent role in our analysis. Given  $i, j \in N$  with

$i \neq j$  set:

$$\alpha(i, j) = \sup \left\{ \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} : \right. \\ \left. a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i) \right\}$$

$$\beta(i, j) = \inf \left\{ \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} : \right. \\ \left. a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) \leq u_j(\tilde{\mathbf{a}}^i), \rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) < \rho(b|\tilde{\mathbf{a}}^i) \right\}$$

Note that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)$  is the gain or loss to player  $j$  from deviating from  $i$ 's preferred action profile  $\tilde{\mathbf{a}}^i$  and  $\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)$  is the increase or decrease in the probability that the bad signal occurs (equivalently, the decrease or increase in the probability that the good signal occurs) following the same deviation. In the definition of  $\alpha(i, j)$  we consider only deviations that are strictly profitable; by assumption, such deviations exist and strictly increase the probability that the bad signal occurs. In view of Assumption 3,  $\alpha(i, j)$  is strictly positive. In the definition of  $\beta(i, j)$  we consider only deviations that are unprofitable *and* strictly decrease the probability that the bad signal occurs, so  $\beta(i, j)$  is the infimum of non-negative numbers and so is necessarily  $+\infty$  (if the infimum is over the empty set) or finite and non-negative.

To gain some intuition, think about how player  $j$  could gain by deviating from  $\tilde{\mathbf{a}}^i$ . On the one hand,  $j$  could gain by deviating to an action that *increases* its current payoff. By assumption, such a deviation will *increase* the probability of a bad signal; assuming that a bad signal leads to a lower continuation utility, whether such a deviation will be profitable will depend on the current gain and on the change in probability;  $\alpha(i, j)$  represents a measure of net profitability from such deviations. On the other hand, player  $j$  could also gain by deviating to an action that *decreases* its current payoff but also *decreases* the probability of a bad signal, and hence leads to a higher continuation utility.  $\beta(i, j)$  represents a measure of net profitability from such deviations.

The measures  $\alpha, \beta$  yield inequalities that must be satisfied in order that there be *any* efficient PPE. Here and throughout, we write  $\text{int } V$  for the interior of  $V$  relative to the hyperplane spanned by  $\{\tilde{v}^i\}$ .

**Proposition 1** Fix  $\delta \in (0, 1)$ . If  $\mathcal{E}(\delta) \cap \text{int } V \neq \emptyset$  then

$$\alpha(i, j) \leq \beta(i, j)$$

for every  $i, j \in N, j \neq i$ .

**Proposition 2** Fix  $\delta \in (0, 1)$ . If  $\mathcal{E}(\delta) \cap \text{int } V \neq \emptyset$  then

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)]$$

for every  $i \in N$  and for all  $a_i \in A_i$ .



The import of Propositions 1 and 2 is that if any of these inequalities fail then efficient payoff vectors with some degree of sharing can *never* be achieved in PPE, no matter what the discount factor is.<sup>11</sup>

We need two further pieces of notation. define

$$\delta^* \triangleq \left( 1 + \frac{1 - \sum_i \lambda_i \bar{v}_i}{\sum_i \left[ \lambda_i \tilde{v}_i^i + \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \right] - 1} \right)^{-1}$$

For each  $i$ , set

$$\bar{v}_i = \max_{j \neq i} \left( \tilde{v}_i^j + \alpha(j, i) [1 - \rho(b|\tilde{\mathbf{a}}^j)] \right)$$

A straightforward but messy computation shows that  $\bar{v}_i$  is at least player  $i$ 's minmax payoff.

**Theorem 1** *Fix  $v \in \text{int}V$ . If*

- (i) for all  $i, j \in N$ ,  $i \neq j$ :  $\alpha(i, j) \leq \beta(i, j)$
- (ii) for all  $i \in N$ ,  $a_i \in A_i$ :

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)]$$

- (iii) for all  $i \in N$ :  $v_i > \bar{v}_i$
- (iv)  $\delta \geq \delta^*$

then  $v$  can be supported in a PPE of  $G^\infty(\delta)$ .<sup>12,13</sup>

The proof of Theorem 1 is explicitly constructive: we provide a simple explicit algorithm that computes a PPE strategy profile that achieves  $v$ . Given the various parameters of the environment (game payoffs, information structure, discount factor) and the target vector  $v$ , the algorithm takes as input in

<sup>11</sup> Proposition 1 might seem somewhat mysterious:  $\alpha$  is a measure of the current gain to deviation and  $\beta$  is a measure of the future gain to deviation; there seems no obvious reason why PPE should necessitate any particular relationship between  $\alpha$  and  $\beta$ . As the proof will show, this relationship arises from the efficiency of payoffs in  $V$  and the assumption that there are only two signals. Taken together, these enable us to identify a crucial quantity (a weighted difference of continuation values) that, at any PPE, must lie (weakly) above  $\alpha$  and (weakly) below  $\beta$ ; in particular it must be the case that  $\alpha$  lies weakly below  $\beta$ .

<sup>12</sup> As we have noted,  $\bar{v}_i$  is at least player  $i$ 's minmax payoff, so (iii) implies that  $v$  dominates the minmax vector, which is of course the familiar necessary condition for achievability in any equilibrium.

<sup>13</sup> Again, we write  $\text{int}V$  for the interior of  $V$  relative to the hyperplane spanned by  $\{\tilde{v}^i\}$ .

**Table 3** The algorithm used by each player.

---

<b>Input:</b> The current continuation payoff $v(t) \in V_\mu$
For each $j$
Calculate the indicator $d_j(v(t))$
Find the player $i$ with largest indicator (if a tie, choose largest $i$ )
$i = \max_j \{\arg \max_{j \in N} d_j(v(t))\}$
Player $i$ is active; chooses action $\tilde{\mathbf{a}}_i^i$
Players $j \neq i$ are inactive; choose action $\tilde{\mathbf{a}}_j^i$
Update $v(t+1)$ as follows:
<b>if</b> $y^t = g$ <b>then</b>
$v_i(t+1) = \tilde{v}_i^i + (1/\delta)(v_i(t) - \tilde{v}_i^i) - (1/\delta - 1)(1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b \tilde{\mathbf{a}}^i)$
$v_j(t+1) = \tilde{v}_j^i + (1/\delta)(v_j(t) - \tilde{v}_j^i) + (1/\delta - 1)\alpha(i, j)\rho(b \tilde{\mathbf{a}}^i)$
for all $j \neq i$
<b>if</b> $y^t = b$ <b>then</b>
$v_i(t+1) = \tilde{v}_i^i + (1/\delta)(v_i(t) - \tilde{v}_i^i) + (1/\delta - 1)(1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(g \tilde{\mathbf{a}}^i)$
$v_j(t+1) = \tilde{v}_j^i + (1/\delta)(v_j(t) - \tilde{v}_j^i) - (1/\delta - 1)\alpha(i, j)\rho(g \tilde{\mathbf{a}}^i)$
for all $j \neq i$

---

period  $t$  a current ‘continuation’ vector  $v(t)$  and computes, for each player  $j$ , a score  $d_j(v(t))$  defined as follows:

$$d_j(v(t)) = \frac{\lambda_j[v_j(t) - \bar{v}_j]}{\lambda_j[\tilde{v}_j^j - v_j(t)] + \sum_{k \neq j} \lambda_k \alpha(j, k) \rho(b|\tilde{\mathbf{a}}^j)}.$$

(We initialize the algorithm by setting  $v(0) = v$ .) Note that each player can compute every score  $d_j$  from the current continuation vector  $v(t)$  and the various parameters. Having computed the entire score vector,  $d(v(t))$ , the algorithm finds the player  $i^*$  whose score  $d_j(v(t))$  is greatest. (In case of ties, we arbitrarily choose the player with the largest index.) The current action profile is  $i^*$ ’s preferred action profile  $\tilde{\mathbf{a}}^{i^*}$ . The algorithm then computes the next period continuation vector as a function of which signal in  $Y$  is realized.

Some intuition may be useful. Each player’s score  $d_j(v(t))$  represents the distance from that player’s current cumulative payoff to that player’s target payoff, with appropriate weights. The player whose score is greatest is therefore the player who is ‘most deserving’ of play in the current period following its preferred action profile. The ‘appropriate weights’ reflect both the payoffs in the stage game and the monitoring structure, and are chosen to yield a strategy profile that is a PPE and also achieves the desired target payoff vector.

## 5 Self-Generating Sets

Our approach to Theorem 1 is to identify a class of sets that are natural candidates for self-generating sets in the sense of APS, show that the Conditions we have are sufficient for these sets to be self-generating, and then show that

the desired target vector lies in one of these sets. In fact, we show that the Conditions are also necessary for these sets to be self-generating; since this seems of some interest in itself, we present it as a separate Theorem.

We begin by recalling some notions from APS. Fix a subset  $W \subset \text{co}(U)$  and a *target payoff*  $v \in \text{co}(U)$ . The target payoff  $v$  can be *decomposed with respect to* the set  $W$  and the discount factor  $\delta < 1$  if there exist an action profile  $\mathbf{a} \in \mathbf{A}$  and continuation payoffs  $\gamma : Y \rightarrow W$  such that

- $v$  is the (weighted) average of current and continuation payoffs when players follow  $\mathbf{a}$

$$v = (1 - \delta)u(\mathbf{a}) + \delta \sum_{y \in Y} \rho(y|\mathbf{a})\gamma(y)$$

- continuation payoffs provide no incentive to deviate: for each  $j$  and each  $a_j \in A_j$

$$v_j \geq (1 - \delta)u_j(a_j, \mathbf{a}_{-j}) + \delta \sum_{y \in Y} \rho(y|a_j, \mathbf{a}_{-j})\gamma_j(y)$$

Write  $\mathcal{B}(W, \delta)$  for the set of target payoffs  $v \in \text{co}(U)$  that can be decomposed with respect to  $W$  for the discount factor  $\delta$ . The set  $W$  is *self-generating* if  $W \subset \mathcal{B}(W, \delta)$ ; i.e., every target vector in  $W$  can be decomposed with respect to  $W$ .

By assumption,  $V$  lies in the bounding hyperplane  $H$ . Hence if we write  $v \in V$  as a convex combination  $v = ax + (1-a)x'$  with  $x, x' \in \text{co}(U)$  then both  $x, x' \in V$ . In particular, if it is possible to decompose  $v \in V$  with respect to *any* set and *any* discount factor, then the utility  $u(\mathbf{a})$  of the associated action profile  $\mathbf{a}$  and the continuation payoffs must lie in  $V$ , and so the associated action profile  $\mathbf{a}$  must lie in  $\tilde{\mathbf{A}}$ . Because we are interested in efficient payoffs we can therefore restrict our search for self-generating sets to subsets  $W \subset V$ . In order to understand which sets  $W \subset V$  can be self-generating, we need to understand how players might profitably gain from deviating from the current recommended action profile. Because we are interested in subsets  $W \subset V$ , the current recommended action profile will always be  $\tilde{\mathbf{a}}^i$  for some  $i$ , so we need to ask how a player  $j$  might profitably gain from deviating from  $\tilde{\mathbf{a}}^i$ . As we have already noted, when  $i$  is the active player, a profitable deviation for player  $j \neq i$  might occur in one of two ways:  $j$  might gain by choosing an action  $a_j \neq \tilde{\mathbf{a}}_j^i$  that increases  $j$ 's *current* payoff or by choosing an action  $a_j \neq \tilde{\mathbf{a}}_j^i$  that alters the signal distribution in such a way as to increase  $j$ 's *future* payoff. Because  $\tilde{\mathbf{a}}^i$  yields  $i$  its best current payoff, a profitable deviation by  $i$  might occur only by choosing an action that that alters the signal distribution in such a way as to increase  $i$ 's *future* payoff. In all cases, the issue will be the net of the current gain/loss against the future loss/gain.

We focus attention on sets of the form

$$V_\mu = \{v \in V : v_i \geq \mu_i \text{ for each } i\}$$

where  $\mu \in \mathbb{R}^n$ . We assume throughout that  $\mu_i > \max_{j \neq i} \tilde{v}_i^j$  and  $V_\mu \neq \emptyset$ . This guarantees that when  $V_\mu$  is not empty, we have  $V_\mu \subset \text{int } V$ ; see Figure 4.

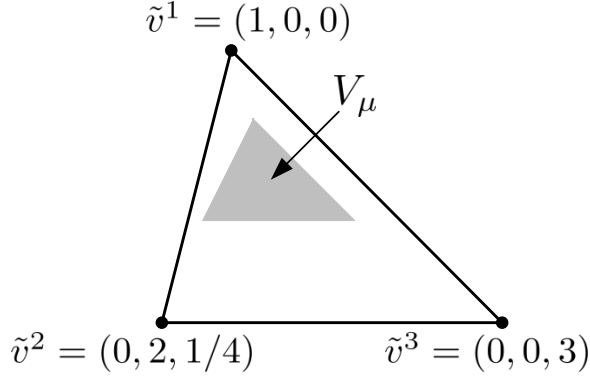


Fig. 4  $\mu = (1/2, 1/2, 1/2)$

The following result shows that the four conditions we have identified (on  $\mu$ , the payoff structure, the information structure  $Y, \rho$  and the discount factor  $\delta$ ) are both *necessary and sufficient* that such a set  $V_\mu$  be self-generating.

**Theorem 2** *Fix the stage game  $G$ , the monitoring structure  $Y, \rho$ , the discount factor  $\delta$  and the vector  $\mu$  with  $\mu_i > \max_{j \neq i} \tilde{v}_i^j$  for all  $i \in N$ . Suppose that  $V_\mu$  has a non-empty interior. In order that the set  $V_\mu$  be self-generating, it is necessary and sufficient that the following four conditions be satisfied.*

- (i) for all  $i, j \in N, i \neq j: \alpha(i, j) \leq \beta(i, j)$
- (ii) for all  $i \in N, a_i \in A_i$ :

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)]$$

- (iii) for all  $i \in N: \mu_i \geq \bar{v}_i$
- (iv) the discount factor  $\delta$  satisfies

$$\delta \geq \delta_\mu \triangleq \left( 1 + \frac{1 - \sum_i \lambda_i \mu_i}{\sum_i \left[ \lambda_i \tilde{v}_i^i + \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \right] - 1} \right)^{-1}$$

One way to contrast our approach with that of FLM is to think about the constraints that need to be satisfied to decompose a given target payoff  $v$  with respect to a given set  $V_\mu$ . By definition we must find a current action profile  $\mathbf{a}$  and continuation payoffs  $\gamma$ . The achievability condition (that  $v$  is the weighted combination of the utility of the current action profile and the expected continuation values) yields a family of linear equalities. The incentive compatibility conditions (that players must be deterred from deviating from

$\mathbf{a}$ ) yields a family of linear inequalities. In the context of FLM, satisfying all these linear inequalities simultaneously requires a large and rich collection of signals so that many different continuation payoffs can be assigned to different deviations. Because we have only two signals, we are only able to choose two continuation payoffs but still must satisfy the same family of inequalities – so our task is much more difficult. It is this difficulty that leads to the Conditions in Theorem 2.

Note that  $\delta_\mu$  is *decreasing* in  $\mu$ . Since Condition 3 puts an absolute lower bound on  $\mu$  and Condition 4 puts an absolute lower bound on  $\delta_\mu$  this means that there is a  $\mu^*$  such that  $V_{\mu^*}$  is the largest self-generating set (of this form) and  $\delta_{\mu^*}$  is the smallest discount factor (for which any set of this form can be self-generating). This may seem puzzling – increasing the discount factor beyond a point makes no difference – but remember that we are providing a characterization of self-generating sets and not of PPE payoffs. However, as we shall see in Theorem 3, for the two-player case, we do obtain a complete characterization of (efficient) PPE payoffs and we demonstrate the same phenomenon.

## 6 Two Players

Theorem 2 provides a complete characterization of self-generating sets that have a special form. If there are only two players then maximal self-generating sets – the set of all PPE – have this form and so it is possible to provide a complete characterization of PPE under the additional assumption that the monitoring structure has full support. We focus on what seems to be the most striking finding: either there are *no efficient PPE outcomes at all for any discount factor*  $\delta < 1$  or there is a discount factor  $\delta^* < 1$  with the property that any target payoff in  $V$  that can be achieved as a PPE for *some*  $\delta$  can already be achieved for *every*  $\delta \geq \delta^*$ .<sup>14</sup>

**Theorem 3** *If  $N = 2$  (two players) and the monitoring structure has full support (i.e.  $0 < \rho(g|\mathbf{a}) < 1$  for each action profile  $\mathbf{a}$ ), then either*

- (i) *no efficient payoff can be supported in a PPE for any discount factor  $\delta < 1$  or*
- (ii) *there exist  $\mu_1^*, \mu_2^*$  and a discount factor  $\delta^* < 1$  such that if  $\delta^* \leq \delta < 1$  then the set of payoff vectors that can be supported in a PPE when the discount factor is  $\delta$  is precisely*

$$E = \{v \in V : v_i \geq \mu_i^* \text{ for } i = 1, 2\}$$

The proof yields explicit (messy) expressions for  $\mu_1^*, \mu_2^*$  and  $\delta^*$ .

<sup>14</sup> The results of Theorem 3 suggest comparison with Proposition 4.11 in an unpublished Working Paper version of FLT. Part 1 of Proposition 4.11 provides sufficient conditions that there be *no PPE*; Theorem 3 is sharper. Part 2 of of Proposition 4.11 assumes that the monitoring structure satisfies “perfect detectability” which seems to require more than two signals, and in any case is not satisfied in our setting.

## 7 Conclusion

This paper contributes to the literature on repeated games with imperfect public monitoring. It makes stronger assumptions about the stage game and the monitoring structure than are common in the literature (the closest comparisons are Fudenberg, Levine, and Maskin (1994) and Fudenberg, Levine, and Takahashi (2007)) and uses those stronger assumptions to make stronger conclusions about efficient PPE. In particular, it proves bounds on the patience players must possess (i.e. on the discount factor) in order that specific efficient outcomes be supportable in PPE and it provides explicit constructions of PPE strategies that support these outcomes.

Clearly there is more to be done in a variety of directions. The monitoring structure has an enormous influence on the structure of efficient PPE (and of PPE more generally). If we view the signal/monitoring structure as the choice made by some agency then (as Example 3 suggests), we might view the interaction as being among  $n + 1$  players: an agency that acts only at the beginning and sets the signal/monitoring structure, which forms part of the “rules” that govern the interaction of the remaining  $n$  players, who interact repeatedly in the stage game. van der Schaar, Xiao, and Zame (2013) indicates a few tentative steps in this direction.

## Appendix

**Proof of Proposition 1** Fix an active player  $i$  and an inactive player  $j$ . Set

$$\begin{aligned} A(i, j) &= \{a_j \in A_j : u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)\} \\ B(i, j) &= \{a_j \in A_j : u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) \leq u_j(\tilde{\mathbf{a}}^i), \rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) < \rho(b|\tilde{\mathbf{a}}^i)\} \end{aligned}$$

If either of  $A(i, j)$  or  $B(i, j)$  is empty then  $\alpha(i, j) \leq \beta(i, j)$  by default, so assume in what follows that neither of  $A(i, j)$ ,  $B(i, j)$  is empty.

Fix a discount factor  $\delta \in (0, 1)$  and let  $\sigma$  be PPE that achieves an efficient payoff. Assume that  $i$  is active following some history:  $\sigma(h) = \tilde{\mathbf{a}}^i$  for some  $h$ . Because  $\sigma$  achieves an efficient payoff, we can decompose the payoff  $v$  following  $h$  as the weighted sum of the current payoff from  $\tilde{\mathbf{a}}^i$  and the continuation payoff assuming that players follow  $\sigma$ ; because  $\sigma$  is a PPE, the incentive compatibility condition for all players  $j$  must obtain. Hence for all  $a_j \in A_j$  we have

$$\begin{aligned} v_j &= (1 - \delta)u_j(\tilde{\mathbf{a}}^i) + \delta \sum_{y \in Y} \rho(y|\tilde{\mathbf{a}}^i)\gamma_j(y) \\ &\geq (1 - \delta)u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) + \delta \sum_{y \in Y} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i)\gamma_j(y), \end{aligned} \quad (1)$$

Substituting probabilities for the good and bad signals yields

$$\begin{aligned} v_j &= (1 - \delta)u_j(\tilde{\mathbf{a}}^i) + \delta [\rho(g|\tilde{\mathbf{a}}^i)\gamma_j(g) + \rho(b|\tilde{\mathbf{a}}^i)\gamma_j(b)] \\ &\geq (1 - \delta)u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) + \delta [\rho(g|a_j, \tilde{\mathbf{a}}_{-j}^i)\gamma_j(g) + \rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i)\gamma_j(b)] \end{aligned}$$

Rearranging yields

$$[\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)] [\gamma_j(g) - \gamma_j(b)] \left[ \frac{\delta}{1-\delta} \right] \geq [u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)]$$

Now suppose  $j \neq i$  is an inactive player. If  $a_j \in A(i, j)$  then  $\rho(y_b^i|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i|\tilde{\mathbf{a}}^i) > 0$  (by Assumption 3) so

$$[\gamma_j(g) - \gamma_j(b)] \left[ \frac{\delta}{1-\delta} \right] \geq \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} \quad (2)$$

If  $a_j \in B(i, j)$  then  $\rho(y_b^i|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i|\tilde{\mathbf{a}}^i) < 0$  (by definition) so

$$[\gamma_j(g) - \gamma_j(b)] \left[ \frac{\delta}{1-\delta} \right] \leq \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} \quad (3)$$

Taking the sup over  $a_j \in A(i, j)$  in (2) and the inf over  $a_j \in B(i, j)$  in (3) yields  $\alpha(i, j) \leq \beta(i, j)$  as desired.

Finally, if  $\mathcal{E}(\delta) \cap \text{int } V \neq \emptyset$ , to achieve any efficient equilibrium payoff in  $\text{int } V$ , every player  $i$  must be active following some history. Hence, we must have  $\alpha(i, j) \leq \beta(i, j)$  for any  $i, j \in N$ ,  $i \neq j$ .  $\square$

**Proof of Proposition 2** As above, we assume  $i$  is active following the history  $h$  and that  $v$  is the payoff following  $h$ . Fix  $a_i \in A_i$ . By definition,  $u_i(\tilde{\mathbf{a}}^i) \geq u_i(a_i, \tilde{\mathbf{a}}_{-i}^i)$ . With respect to probabilities, there are two possibilities. If  $\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) \leq \rho(b|\tilde{\mathbf{a}}^i)$  then we immediately have

$$\tilde{v}_i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)]$$

because the left-hand side is non-negative and the right-hand side is non-positive ( $\alpha(i, j)$  is positive due to Assumption 3). If  $\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) > \rho(b|\tilde{\mathbf{a}}^i)$  we proceed as follows.

We begin with (1) but now we apply it to the active user  $i$ , so that for all  $a_i \in A_i$  we have

$$\begin{aligned} v_i &= (1-\delta)u_i(\tilde{\mathbf{a}}^i) + \delta[\rho(g|\tilde{\mathbf{a}}^i)\gamma_i(g) + \rho(b|\tilde{\mathbf{a}}^i)\gamma_i(b)] \\ &\geq (1-\delta)u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) + \delta[(\rho(g|a_i, \tilde{\mathbf{a}}_{-i}^i)\gamma_i(g) + \rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i)\gamma_i(b)] \end{aligned}$$

Rearranging yields

$$\gamma_i(g) - \gamma_i(b) \geq \left[ \frac{1-\delta}{\delta} \right] \left[ \frac{u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) - u_i(\tilde{\mathbf{a}}^i)}{\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)} \right]$$

Because continuation payoffs are in  $V$ , which lies in the hyperplane  $H$ , the continuation payoffs for the active user can be expressed in terms of the continuation payoffs for the inactive users as

$$\gamma_i(y) = \frac{1}{\lambda_i} \left[ 1 - \sum_{j \neq i} \lambda_j \gamma_j(y) \right]$$

Hence

$$\gamma_i(g) - \gamma_i(b) = -\frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j [\gamma_j(g) - \gamma_j(b)]$$

Applying the incentive compatibility constraints for the inactive users implies that for each  $a_j \in A(i, j)$  we have

$$\gamma_j(g) - \gamma_j(b) \geq \left[ \frac{1 - \delta}{\delta} \right] \left[ \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} \right]$$

In particular

$$\gamma_j(g) - \gamma_j(b) \geq \left[ \frac{1 - \delta}{\delta} \right] \alpha(i, j)$$

and hence

$$\gamma_i(g) - \gamma_i(b) \leq -\frac{1}{\lambda_i} \left[ \frac{1 - \delta}{\delta} \right] \left[ \sum_{j \neq i} \lambda_j \alpha(i, j) \right]$$

Putting these all together, canceling the factor  $[1 - \delta]/\delta$  and remembering that we are in the case  $\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) > \rho(b|\tilde{\mathbf{a}}^i)$  yields

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(b|a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(b|\tilde{\mathbf{a}}^i)]$$

which is the desired result. Again, if  $\mathcal{E}(\delta) \cap \text{int } V \neq \emptyset$ , every player  $i$  must be active after some history to achieve some PPE payoff in  $\text{int } V$ . Hence, the above inequality must hold for every  $i \in N$ .  $\square$

**Proof of Theorem 1** Theorem 1 is a straightforward consequence of Theorem 2. Specifically, for any  $v \in \text{int } V$  that satisfies  $v_i > \bar{v}_i$  for all  $i \in N$ , we define the set

$$V_{\bar{v}} = \{v \in V : v_i \geq \bar{v}_i \text{ for each } i\}.$$

Clearly,  $V_{\bar{v}}$  contains  $v$ . Since  $v_i > \bar{v}_i$  for all  $i \in N$ ,  $V_{\bar{v}}$  has a non-empty interior. From the definition that  $\bar{v}_i = \max_{j \neq i} \left( \tilde{v}_i^j + \alpha(j, i)[1 - \rho(b|\tilde{\mathbf{a}}^j)] \right)$ , we know that  $\bar{v}_i > \max_{j \neq i} \tilde{v}_i^j$ , because  $\alpha(j, i) > 0$  (due to Assumption 3) and  $1 - \rho(b|\tilde{\mathbf{a}}^j) > 0$  (due to Assumption 3, we have  $1 - \rho(b|\tilde{\mathbf{a}}^j) = \rho(g|\tilde{\mathbf{a}}^j) > \rho(g|a_i, \tilde{\mathbf{a}}_{-i}^j) \geq 0$ , where  $a_i$  is the action such that  $u_i(a_i, \tilde{\mathbf{a}}_{-i}^j) > u_i(\tilde{\mathbf{a}}^j)$ ). Hence,  $V_{\bar{v}}$  must be in the interior of  $V$ .

When the sufficient conditions in Theorem 1 hold, Theorem 2 guarantees that the set  $V_{\bar{v}}$  is a self-generating set for any discount factor  $\delta \geq \delta^*$ . Hence, the target payoff  $v \in V_{\bar{v}}$  can be achieved in a PPE for any discount factor  $\delta \geq \delta^*$ .  $\square$

**Proof of Theorem 2** We first prove that the four conditions are necessary. Assume that  $V_\mu$  is a self-generating set; we verify Conditions (i)-(iv) in turn. Since we focus on  $\mu$  that satisfies  $\mu_i > \max_{j \neq i} \tilde{v}_i^j$ , we have  $V_\mu \subset \text{int } V$ . Since



$V_\mu \subseteq \mathcal{E}(\delta)$ ,  $\mathcal{E}(\delta) \cap \text{int } V$  must be non-empty. Hence, Propositions 1 and 2 yield Conditions (i) and (ii).

Now we derive Condition (iii). To do this, we suppose  $i$  is active and examine the decomposition of the inactive player  $j$ 's payoff in greater detail. Because  $\mu_j > \tilde{v}_j^i$  and  $v_j \geq \mu_j$  for every  $v \in V_\mu$  we certainly have  $v_j > \tilde{v}_j^i$ . We can write  $j$ 's incentive compatibility condition as

$$\begin{aligned} v_j &= (1 - \delta) \cdot \tilde{v}_j^i + \delta \cdot \sum_{y \in Y} \rho(y|\tilde{\mathbf{a}}^i) \cdot \gamma_j(y) \\ &\geq (1 - \delta) \cdot u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) + \delta \cdot \sum_{y \in Y} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(y). \end{aligned} \quad (4)$$

From the equality constraint in (4), we can solve for the discount factor  $\delta$  as

$$\delta = \frac{v_j - \tilde{v}_j^i}{\sum_{y \in Y} \gamma_j(y) \rho(y|\tilde{\mathbf{a}}^i) - \tilde{v}_j^i}$$

(Note that the denominator can never be zero and the above equation is well defined, because  $v_j > \tilde{v}_j^i$  implies that  $\sum_{y \in Y} \gamma_j(y) \rho(y|\tilde{\mathbf{a}}^i) > \tilde{v}_j^i$ .) We can then eliminate the discount factor  $\delta$  in the inequality of (4). Since  $v_j > \tilde{v}_j^i$ , we can obtain equivalent inequalities, depending on whether  $a_j$  is a profitable or unprofitable current deviation):

– If  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > \tilde{v}_j^i$  then

$$\begin{aligned} v_j &\leq \sum_{y \in Y} \gamma_j(y) \left[ \left( 1 - \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right) \rho(y|\tilde{\mathbf{a}}^i) \right. \\ &\quad \left. + \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \right] \end{aligned} \quad (5)$$

– If  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) < \tilde{v}_j^i$  then

$$\begin{aligned} v_j &\geq \sum_{y \in Y} \gamma_j(y) \left[ \left( 1 - \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right) \rho(y|\tilde{\mathbf{a}}^i) \right. \\ &\quad \left. + \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \right] \end{aligned} \quad (6)$$

For notational convenience, write the coefficient of  $\gamma_j(g)$  in the above inequalities as

$$\begin{aligned} c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) &\triangleq \left(1 - \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i}\right) \rho(g|\tilde{\mathbf{a}}^i) \\ &\quad + \left(\frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i}\right) \rho(g|a_j, \tilde{\mathbf{a}}_{-j}^i) \\ &= \rho(g|\tilde{\mathbf{a}}^i) + (v_j - \tilde{v}_j^i) \left(\frac{\rho(g|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(g|\tilde{\mathbf{a}}^i)}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i}\right) \\ &= \rho(g|\tilde{\mathbf{a}}^i) - (v_j - \tilde{v}_j^i) \left(\frac{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i}\right) \end{aligned}$$

According to (5), if  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > \tilde{v}_j^i$  then

$$c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(g) + [1 - c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i)] \gamma_j(b) \leq v_j \quad (7)$$

Since  $\gamma_j(g) > \gamma_j(b)$ , this is true if and only if

$$\kappa_{ij}^+ \cdot \gamma_j(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j(b) \leq v_j, \quad (8)$$

where  $\kappa_{ij}^+ \triangleq \sup\{c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) : a_j \in A_j : u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > \tilde{v}_j^i\}$ . (Fulfilling the inequalities (7) for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$  is equivalent to fulfilling the single inequality (8). If (8) is satisfied, then the inequalities (7) are satisfied for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$  because  $\gamma_j(g) > \gamma_j(b)$  and  $\kappa_{ij}^+ \geq c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i)$  for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$ . Conversely, if the inequalities (7) are satisfied for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$  and (8) were violated, so that  $\kappa_{ij}^+ \cdot \gamma_j(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j(b) > v_j$ , then we can find a  $\kappa'_{ij} < \kappa_{ij}^+$  such that  $\kappa'_{ij} \cdot \gamma_j(g) + (1 - \kappa'_{ij}) \cdot \gamma_j(b) > v_j$ . Based on the definition of the supremum, there exists at least a  $a'_j$  such that  $u_j(a'_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$  and  $c_{ij}(a'_j, \tilde{\mathbf{a}}_{-j}^i) > \kappa'_{ij}$ , which means that  $c_{ij}(a'_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(g) + (1 - c_{ij}(a'_j, \tilde{\mathbf{a}}_{-j}^i)) \cdot \gamma_j(b) > v_j$ . This contradicts the fact that the inequalities (8) are fulfilled for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$ .)

Similarly, according to (6), for all  $a_j$  such that  $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) < \tilde{v}_j^i$ , we must have

$$c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) \gamma_j(g) + [1 - c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i)] \gamma_j(b) \geq v_j.$$

Since  $\gamma_j(g) > \gamma_j(b)$ , the above requirement is fulfilled if and only if

$$\kappa_{ij}^- \cdot \gamma_j(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j(b) \geq v_j,$$

where  $\kappa_{ij}^- \triangleq \inf\{c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) : a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) < \tilde{v}_j^i\}$ . Hence, the decomposition (4) for user  $j \neq i$  can be simplified as:

$$\begin{aligned} \rho(g|\tilde{\mathbf{a}}^i) \cdot \gamma_j(g) + [1 - \rho(g|\tilde{\mathbf{a}}^i)] \gamma_j(b) &= \tilde{v}_j^i + \frac{v_j - \tilde{v}_j^i}{\delta} \\ \kappa_{ij}^+ \gamma_j(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j(b) &\leq v_j \\ \kappa_{ij}^- \gamma_j(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j(b) &\geq v_j \end{aligned} \quad (9)$$

Keep in mind that the various continuation values  $\gamma$  and the expressions  $\kappa_{ij}^+, \kappa_{ij}^-$  depend on  $v_j$ ; where necessary we write the dependence explicitly. Note that there could be many  $\gamma_j(g)$  and  $\gamma_j(b)$  that satisfy (9). For a given discount factor  $\delta$ , we call all the continuation payoffs that satisfy (9) *feasible* – but whether particular continuation values lie in  $V_\mu$  depends on the discount factor.

We assert that  $\kappa_{ij}^+(\mu_j) \leq 0$  for all  $i \in N$  and for all  $j \neq i$ . To see this, we look at a payoff profile  $\hat{v}^i$  defined as

$$\hat{v}_j^i = \begin{cases} \mu_j & \text{if } j \neq i \\ \frac{1}{\lambda_i} \left(1 - \sum_{k \neq i} \lambda_k \mu_k\right) & \text{if } j = i \end{cases}.$$

We can prove that the payoff profile  $\hat{v}^i$  indeed lies in  $V_\mu$ . In fact, the desired payoff profile  $\hat{v}^i$  is the maximizer of the following optimization problem:  $\max_{v \in V_\mu} v_i$ . Since  $V_\mu$  is compact, the solution to the optimization problem  $\max_{v \in V_\mu} v_i$  exists. Suppose that the solution is  $v^* \neq \hat{v}^i$ , namely there exists a  $j \neq i$  such that  $v_j^* > \mu_j$ . Then we can define a vector  $v'$  with a slightly lower payoff for player  $j$  and a slightly larger payoff for player  $i$ , namely:

$$v'_j = v_j^* - \epsilon, \quad v'_i = v_i^* + \frac{\lambda_j}{\lambda_i} \epsilon, \quad v'_k = v_k^*, \quad \forall k \neq i, j.$$

Clearly, we have  $v' \in H$ . Since  $V_\mu \subset \text{int}V$ , we can find a small enough  $\epsilon$  such that  $v' \in V$  and that  $v'_j \geq \mu_j$ . Hence,  $v'$  is in  $V_\mu$ , and has a higher payoff for player  $i$ . This is contradictory to the fact that  $v^*$  is the solution to the problem  $\max_{v \in V_\mu} v_i$ . Hence, the maximizer of  $\max_{v \in V_\mu} v_i$  must be  $\hat{v}^i$ . Therefore, we must have  $\hat{v}^i \in V_\mu$ .

Since  $\hat{v}^i \in V_\mu$ , the payoff profile  $\hat{v}^i$  must be decomposable. Observe that in the payoff profile  $\hat{v}^i$ , all the players  $j \neq i$  get the lowest possible payoffs  $\mu_j$  in  $V_\mu$ , and player  $i$  gets the highest possible payoff in  $V_\mu$ . As a result,  $\hat{v}^i$  is necessarily decomposed by  $\tilde{\alpha}^i$ . We look at the following constraint for player  $j \neq i$  in (9):

$$\kappa_{ij}^+ \gamma_j(g) + (1 - \kappa_{ij}^+) \gamma_j(b) \leq \mu_j.$$

Suppose that  $\kappa_{ij}^+(\mu_j) > 0$ . Since player  $j$  has a currently profitable deviation from  $\tilde{\alpha}^i$ , we must set  $\gamma_j(g) > \gamma_j(b)$ . Then to satisfy the above inequality, we must have  $\gamma_j(b) < \mu_j$ . In other words, when  $\kappa_{ij}^+(\mu_j) > 0$ , all the feasible continuation payoffs of player  $j$  must be outside  $V_\mu$ . This contradicts the fact that  $V_\mu$  is self-generating so the assertion follows.

The definition of  $\kappa_{ij}^+(\mu_j)$  and the fact that  $\kappa_{ij}^+(\mu_j) \leq 0$  entail that

$$\begin{aligned}
\kappa_{ij}^+(\mu_j) &= \sup_{a_j \in A(i,j)} \left\{ \rho(g|\tilde{\mathbf{a}}^i) - (\mu_j - \tilde{v}_j^i) \left[ \frac{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right] \right\} \\
&= \rho(g|\tilde{\mathbf{a}}^i) - (\mu_j - \tilde{v}_j^i) \inf_{a_j \in A(i,j)} \left[ \frac{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right] \\
&= \rho(g|\tilde{\mathbf{a}}^i) - (\mu_j - \tilde{v}_j^i) \left[ \frac{1}{\sup_{a_j \in A(i,j)} \left( \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i}{\rho(b|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(b|\tilde{\mathbf{a}}^i)} \right)} \right] \\
&= \rho(g|\tilde{\mathbf{a}}^i) - (\mu_j - \tilde{v}_j^i) \left[ \frac{1}{\alpha(i,j)} \right] \\
&\leq 0
\end{aligned}$$

This provides a lower bound on  $\mu_j$ :

$$\mu_j \geq \tilde{v}_j^i + \alpha(i,j)\rho(g|\tilde{\mathbf{a}}^i) = \tilde{v}_j^i + \alpha(i,j)[1 - \rho(b|\tilde{\mathbf{a}}^i)]$$

This bound must hold for every  $i \in N$  and every  $j \neq i$ . Hence, we have

$$\mu_j \geq \max_{i \neq j} (\tilde{v}_j^i + \alpha(i,j)[1 - \rho(b|\tilde{\mathbf{a}}^i)])$$

which is Condition (iii).

Now we derive Condition (iv), the necessary condition on the discount factor. The minimum discount factor  $\delta_\mu$  required for  $V_\mu$  to be a self-generating set solves the optimization problem

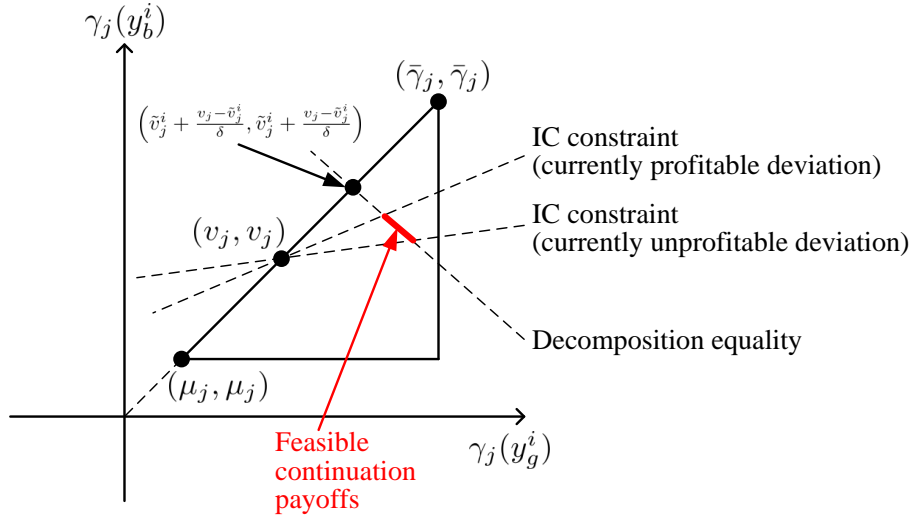
$$\delta_\mu = \max_{v \in V_\mu} \delta \quad \text{subject to } v \in \mathcal{B}(V_\mu, \delta)$$

where  $\mathcal{B}(V_\mu, \delta)$  is the set of payoff profiles that can be decomposed on  $V_\mu$  under discount factor  $\delta$ . Since  $\mathcal{B}(V_\mu; \delta) = \cup_{i \in N} \mathcal{B}(V_\mu, \delta, \tilde{\mathbf{a}}^i)$ , where  $\mathcal{B}(V_\mu, \delta, \tilde{\mathbf{a}}^i)$  is the set of payoff profiles that can be decomposed on  $V_\mu$  by  $\tilde{\mathbf{a}}^i$  under discount factor  $\delta$ , the above optimization problem can be reformulated as

$$\delta_\mu = \max_{v \in V_\mu} \min_{i \in N} \delta \quad \text{subject to } v \in \mathcal{B}(V_\mu, \delta, \tilde{\mathbf{a}}^i). \quad (10)$$

To solve the optimization problem (10), we explicitly express the constraint  $v \in \mathcal{B}(V_\mu, \delta, \tilde{\mathbf{a}}^i)$  using the results derived above.

Some intuition may be useful. Suppose that  $i$  is active and  $j$  is an inactive player. Recall that player  $j$ 's feasible  $\gamma_j(g)$  and  $\gamma_j(b)$  must satisfy (9). There are many  $\gamma_j(g)$  and  $\gamma_j(b)$  that satisfy (9). In Fig. 5, we show the feasible continuation payoffs that satisfy (9) when  $\kappa_{ij}^+(v_j) \leq 0$ . We can see that all the continuation payoffs on the heavy line segment are feasible. The line segment is on the line that represents the decomposition equality  $\rho(g|\tilde{\mathbf{a}}^i) \cdot \gamma_j(g) + (1 - \rho(g|\tilde{\mathbf{a}}^i)) \cdot \gamma_j(b) = \tilde{v}_j^i + \frac{v_j - \tilde{v}_j^i}{\delta}$ , and is bounded by the IC constraint on currently



**Fig. 5** Illustrations of the feasible continuation payoffs when  $\kappa_{ij}^+ \leq 0$ .  $\bar{\gamma}_j = \frac{1}{\lambda_j} (1 - \sum_{k \neq j} \lambda_k \mu_k)$ .

profitable deviations  $\kappa_{ij}^+ \cdot \gamma_j(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j(b) \leq v_j$  and the IC constraint on currently unprofitable deviations  $\kappa_{ij}^- \cdot \gamma_j(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j(b) \geq v_j$ . Among all the feasible continuation payoffs, denoted  $\gamma'(y)$ , we choose the one, denoted  $\gamma^*(y)$ , such that for all  $j \neq i$ ,  $\gamma_j^*(g)$  and  $\gamma_j^*(b)$  make the IC constraint on currently profitable deviations in (9) binding. This is because under the same discount factor  $\delta$ , if there is any feasible continuation payoff  $\gamma'(y)$  in the self-generating set, the one that makes the IC constraint on currently profitable deviations binding is also in the self-generating set. The reason is that, as can be seen from Fig. 5, the continuation payoff  $\gamma_j^*(y)$  that makes the IC constraint binding has the smallest  $\gamma_j^*(g) = \min \gamma_j'(g)$  and the largest  $\gamma_j^*(b) = \max \gamma_j'(b)$ . Formally we establish the following Lemma.

**Lemma 1** Fix a payoff profile  $v$  and a discount factor  $\delta$ . Suppose that  $v$  is decomposed by  $\tilde{\mathbf{a}}^i$ . If there are any feasible continuation payoffs  $\gamma'(g) \in V_\mu$  and  $\gamma'(b) \in V_\mu$  that satisfy (9) for all  $j \neq i$ , then there exist feasible continuation payoffs  $\gamma^*(g) \in V_\mu$  and  $\gamma^*(b) \in V_\mu$  such that the IC constraint on currently profitable deviations in (9) is binding for all  $j \neq i$ .

*Proof* Given feasible continuation payoffs  $\gamma'(g) \in V_\mu$  and  $\gamma'(b) \in V_\mu$ , we construct  $\gamma^*(g) \in V_\mu$  and  $\gamma^*(b) \in V_\mu$  that are feasible and make the IC constraint on currently profitable deviations in (9) binding for all  $j \neq i$ .

First, define  $\gamma_j^*(g)$  and  $\gamma_j^*(b)$ ,  $\forall j \neq i$  as the solutions to the decomposition equality and the binding IC constraint on currently profitable deviations in

(9):

$$\begin{aligned}\rho(g|\tilde{\mathbf{a}}^i) \cdot \gamma_j(g) + [1 - \rho(g|\tilde{\mathbf{a}}^i)]\gamma_j(b) &= \tilde{v}_j^i + \frac{v_j - \tilde{v}_j^i}{\delta} \\ \kappa_{ij}^+ \gamma_j(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j(b) &= v_j\end{aligned}$$

We have shown that  $\kappa_{ij}^+ \leq 0$ . Hence,  $\gamma_j^*(g)$  and  $\gamma_j^*(b)$  exist and are unique. In addition, we have

$$\begin{aligned}& [\rho(g|\tilde{\mathbf{a}}^i) \cdot \gamma_j^*(g) + [1 - \rho(g|\tilde{\mathbf{a}}^i)]\gamma_j^*(b)] - [\kappa_{ij}^+ \gamma_j^*(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j^*(b)] \\ &= [\rho(g|\tilde{\mathbf{a}}^i) - \kappa_{ij}^+] \cdot [\gamma_j^*(g) - \gamma_j^*(b)] \\ &= (v_j - \tilde{v}_j^i) \cdot \left( \frac{1}{\delta} - 1 \right) > 0 \\ &\Rightarrow [\rho(g|\tilde{\mathbf{a}}^i) - \kappa_{ij}^+] \cdot [\gamma_j^*(g) - \gamma_j^*(b)] > 0 \\ &\Rightarrow \gamma_j^*(g) > \gamma_j^*(b)\end{aligned}$$

Second, we show that  $\gamma_j^*(g)$  and  $\gamma_j^*(b)$  must satisfy the IC constraint on currently unprofitable deviations in (9):

$$\kappa_{ij}^- \gamma_j(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j(b) \geq v_j$$

Since there exist feasible  $\gamma_j'(g)$  and  $\gamma_j'(b)$ , and since we have shown that  $\gamma_j'(g) > \gamma_j'(b)$ , we have

$$\begin{aligned}& \kappa_{ij}^- \gamma_j'(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j'(b) \geq \kappa_{ij}^+ \gamma_j'(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j'(b) \\ &\Rightarrow [\kappa_{ij}^- - \kappa_{ij}^+] \cdot [\gamma_j'(g) - \gamma_j'(b)] \geq 0 \\ &\Rightarrow \kappa_{ij}^- \geq \kappa_{ij}^+\end{aligned}$$

Hence, we must have

$$\kappa_{ij}^- \gamma_j^*(g) + (1 - \kappa_{ij}^-) \cdot \gamma_j^*(b) \geq \kappa_{ij}^+ \gamma_j^*(g) + (1 - \kappa_{ij}^+) \cdot \gamma_j^*(b) = v_j$$

Finally, we show that  $\gamma^*(y) \in V_\mu$ . For this, we need to prove that  $\gamma_j^*(g) \leq \gamma_j'(g)$  and  $\gamma_j^*(b) \geq \gamma_j'(b)$ . We prove this by contradiction. Suppose that there exist  $\gamma_j'(g)$  and  $\gamma_j'(b)$  that satisfy (9) and  $\gamma_j'(g) = \gamma_j^*(g) - \zeta$  with  $\zeta > 0$ . Based on the decomposition equality, we have

$$\gamma_j'(b) = \gamma_j^*(b) + \left( \frac{\rho(g|\tilde{\mathbf{a}}^i)}{1 - \rho(g|\tilde{\mathbf{a}}^i)} \right) \zeta$$

We can see that the IC constraint on currently profitable deviations is violated:

$$\begin{aligned}& \kappa_{ij}^+ \gamma_j'(g) + (1 - \kappa_{ij}^+) \gamma_j'(b) \\ &= \kappa_{ij}^+ \gamma_j^*(g) + (1 - \kappa_{ij}^+) \gamma_j^*(b) + \left[ -\kappa_{ij}^+ \zeta + (1 - \kappa_{ij}^+) \left( \frac{\rho(g|\tilde{\mathbf{a}}^i)}{1 - \rho(g|\tilde{\mathbf{a}}^i)} \right) \zeta \right] \\ &= v_j + (1 - \kappa_{ij}^+) \left[ \frac{\rho(g|\tilde{\mathbf{a}}^i)}{1 - \rho(g|\tilde{\mathbf{a}}^i)} - \frac{\kappa_{ij}^+}{1 - \kappa_{ij}^+} \right] \zeta \\ &> v_j\end{aligned}$$

where the last inequality results from  $\kappa_{ij}^+ \leq 0$ . Hence, we have  $\gamma_j^*(g) \leq \gamma_j'(g)$  and  $\gamma_j^*(b) \geq \gamma_j'(b)$  for all  $\gamma_j'(g)$  and  $\gamma_j'(b)$  that satisfy (9).

Now we can prove that  $\gamma^*(y) \in V_\mu$ . For this, we need to show that  $\gamma_j^*(g) \geq \mu_j$  and  $\gamma_j^*(b) \geq \mu_j$  for all  $j \in N$ . For  $j \neq i$ , we have  $\gamma_j^*(g) \geq \gamma_j'(g) \geq \mu_j$ . For  $i$ , we have

$$\gamma_i^*(g) = \frac{1}{\lambda_i} \left( 1 - \sum_{j \neq i} \lambda_j \gamma_j^*(g) \right) \geq \frac{1}{\lambda_i} \left( 1 - \sum_{j \neq i} \lambda_j \gamma_j'(g) \right) = \gamma_i'(g) \geq \mu_i$$

This proves the lemma.

Using this Lemma, we can calculate the continuation payoffs of the inactive player  $j \neq i$ :

$$\begin{aligned} \gamma_j(g) &= \frac{(\frac{1}{\delta}(1 - \kappa_{ij}^+) - [1 - \rho(g|\tilde{\mathbf{a}}^i)])v_j - (\frac{1}{\delta} - 1)(1 - \kappa_{ij}^+)\tilde{v}_j^i}{\rho(g|\tilde{\mathbf{a}}^i) - \kappa_{ij}^+} \\ &= \frac{v_j}{\delta} - \left(\frac{1 - \delta}{\delta}\right)\tilde{v}_j^i + \left(\frac{1 - \delta}{\delta}\right)[1 - \rho(g|\tilde{\mathbf{a}}^i)]\alpha(i, j), \\ \gamma_j(b) &= \frac{[\rho(g|\tilde{\mathbf{a}}^i) - \frac{1}{\delta}\kappa_{ij}^+]v_j + (\frac{1}{\delta} - 1)\kappa_{ij}^+\tilde{v}_j^i}{\rho(g|\tilde{\mathbf{a}}^i) - \kappa_{ij}^+} \\ &= \frac{v_j}{\delta} - \left(\frac{1 - \delta}{\delta}\right)\tilde{v}_j^i - \left(\frac{1 - \delta}{\delta}\right)\rho(g|\tilde{\mathbf{a}}^i)\alpha(i, j). \end{aligned}$$

The active player's continuation payoffs can be determined based on the inactive players' continuation payoffs since  $\gamma(y) \in V$ . We calculate the active player  $i$ 's continuation payoffs as

$$\begin{aligned} \gamma_i(g) &= \frac{v_i}{\delta} - \left(\frac{1 - \delta}{\delta}\right)\tilde{v}_i^i - \left(\frac{1 - \delta}{\delta}\right)[1 - \rho(g|\tilde{\mathbf{a}}^i)]\frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j), \\ \gamma_i(b) &= \frac{v_i}{\delta} - \left(\frac{1 - \delta}{\delta}\right)\tilde{v}_i^i + \left(\frac{1 - \delta}{\delta}\right)\rho(g|\tilde{\mathbf{a}}^i)\frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) \end{aligned}$$

Hence, the constraint  $v \in \mathcal{B}(V_\mu, \delta, \tilde{\mathbf{a}}^i)$  on discount factor  $\delta$  is equivalent to

$$\gamma(y) \in V_\mu \text{ for all } y \in Y \Leftrightarrow \gamma_i(y) \geq \mu_i \text{ for all } i \in N, y \in Y$$

Since  $\kappa_{ij}^+(\mu_j) \leq 0$ , we have  $\gamma_j(y) \geq v_j$  for all  $y \in Y$ , which means that  $\gamma_j(y) \geq \mu_j$  for all  $y \in Y$ . Hence, we only need the discount factor to have the property that  $\gamma_i(y) \geq \mu_i$  for all  $y \in Y$ . Since  $\gamma_i(g) < \gamma_i(b)$ , we need  $\gamma_i(g) \geq \mu_i$ , which leads to

$$\delta \geq \frac{1}{1 + \lambda_i(v_i - \mu_i) / \left[ \lambda_i(\tilde{v}_i^i - v_i) + \sum_{j \neq i} \lambda_j \cdot (1 - \rho(g|\tilde{\mathbf{a}}^i))\alpha(i, j) \right]}.$$

Hence, the optimization problem (10) is equivalent to

$$\delta(\mu) = \max_{v \in V_\mu} \min_{i \in \mathcal{N}} x_i(v) \quad (11)$$

where

$$x_i(v) \triangleq \frac{1}{1 + \lambda_i(v_i - \mu_i) / \left( \lambda_i(\tilde{v}_i^i - v_i) + \sum_{j \neq i} \lambda_j [1 - \rho(g|\tilde{\mathbf{a}}^i)] \alpha(i, j) \right)}$$

Since  $x_i(v)$  is decreasing in  $v_i$ , the payoff  $v^*$  that maximizes  $\min_{i \in \mathcal{N}} x_i(v)$  must satisfy  $x_i(v^*) = x_j(v^*)$  for all  $i$  and  $j$ . Now we find the payoff  $v^*$  such that  $x_i(v^*) = x_j(v^*)$  for all  $i$  and  $j$ .

Define

$$z \triangleq \frac{\lambda_i(v_i^* - \mu_i)}{\lambda_i(\tilde{v}_i^i - v_i^*) + \sum_{j \neq i} \lambda_j [1 - \rho(g|\tilde{\mathbf{a}}^i)] \alpha(i, j)}$$

Then we have

$$\lambda_i(1+z)v_i^* = \lambda_i(\mu_i + z\tilde{v}_i^i) - z \sum_{j \neq i} \lambda_j [1 - \rho(g|\tilde{\mathbf{a}}^i)] \alpha(i, j)$$

from which it follows that

$$z = \frac{1 - \sum_i \lambda_i \mu_i}{\sum_i \left( \lambda_i \tilde{v}_i^i + \sum_{j \neq i} \lambda_j [1 - \rho(g|\tilde{\mathbf{a}}^i)] \alpha(i, j) \right) - 1}$$

Hence, the minimum discount factor is  $\delta(\mu) = \frac{1}{1+z}$ ; substituting the definition of  $z$  yields Condition (iv). This completes the proof that these Conditions 1-4 are necessary for  $V_\mu$  to be a self-generating set.

It remains to show that these necessary Conditions are also sufficient. Specifically, we aim to show that under Conditions (i)-(iv), we can decompose each payoff profile  $v \in V_\mu$ .

For convenience, we summarize how we decompose any  $v \in V_\mu$  as follows. We first find the active player  $i$  according to

$$i = \max_j \left\{ \arg \max_{j \in \mathcal{N}} d_j(v) \right\},$$

where

$$d_j(v) = \frac{\lambda_j[v_j - \mu_j]}{\lambda_j[\tilde{v}_j^j - v_j] + \sum_{k \neq j} \lambda_k \alpha(j, k) \rho(b|\tilde{\mathbf{a}}^j)}.$$

Then we assign the continuation payoff vectors  $\gamma(y)$  as follows:

$$\gamma_i(g) = \tilde{v}_i^i + (1/\delta)(v_i(t) - \tilde{v}_i^i) - (1/\delta - 1)(1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i),$$

$$\gamma_j(g) = \tilde{v}_j^j + (1/\delta)(v_j(t) - \tilde{v}_j^j) + (1/\delta - 1) \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i), \forall j \neq i,$$



and

$$\begin{aligned}\gamma_i(b) &= \tilde{v}_i^i + (1/\delta)(v_i(t) - \tilde{v}_i^i) + (1/\delta - 1)(1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(g|\tilde{\mathbf{a}}^i), \\ \gamma_j(b) &= \tilde{v}_j^i + (1/\delta)(v_j(t) - \tilde{v}_j^i) - (1/\delta - 1)\alpha(i, j)\rho(g|\tilde{\mathbf{a}}^i), \forall j \neq i.\end{aligned}$$

We need to verify that under Conditions (i)-(iv), the above continuation payoff vectors  $\gamma(g)$  and  $\gamma(b)$  satisfy 1) the decomposition equalities, 2) the incentive compatibility constraints, and 3) that  $\gamma(g) \in V_\mu$  and  $\gamma(b) \in V_\mu$ .

It is straightforward to check that the decomposition equalities are satisfied. The incentive compatibility constraints for the inactive players  $j$  reduce to Condition (i), and those for the active player  $i$  reduce to Condition (ii).

We proceed to verify that  $\gamma(g) \in V_\mu$  and  $\gamma(b) \in V_\mu$ . It is straightforward to verify that  $\gamma(g) \in V$  and  $\gamma(b) \in V$ . We only need to show  $\gamma_j(g) \geq \mu_j$  and  $\gamma_j(b) \geq \mu_j$  for all  $j \in N$ . Since  $\alpha(i, j) > 0$ , we can observe that  $\gamma_j(g) > \gamma_j(b)$  for all  $j \neq i$  and  $\gamma_i(g) < \gamma_i(b)$ . Hence, it suffices to show  $\gamma_j(b) \geq \mu_j$  for all  $j \neq i$  and  $\gamma_i(g) \geq \mu_i$ .

For any inactive player  $j$ , we have

$$\begin{aligned}\gamma_j(b) &\geq \mu_j \\ \Leftrightarrow \tilde{v}_j^i + (1/\delta)(v_j(t) - \tilde{v}_j^i) - (1/\delta - 1)\alpha(i, j)\rho(g|\tilde{\mathbf{a}}^i) &\geq \mu_j \\ \Leftrightarrow (1/\delta)v_j(t) - \mu_j &\geq (1/\delta - 1)\tilde{v}_j^i + (1/\delta - 1)\alpha(i, j)\rho(g|\tilde{\mathbf{a}}^i) \\ \Leftrightarrow (1/\delta)\mu_j - \mu_j &\geq (1/\delta - 1)\tilde{v}_j^i + (1/\delta - 1)\alpha(i, j)\rho(g|\tilde{\mathbf{a}}^i) \\ \Leftrightarrow \mu_j &\geq \tilde{v}_j^i + \alpha(i, j)\rho(g|\tilde{\mathbf{a}}^i) \\ \Leftrightarrow &\text{Condition (iii)}.\end{aligned}$$

For the active player  $i$ , we have

$$\begin{aligned}
& \gamma_i(g) \geq \mu_i \\
& \Leftrightarrow \tilde{v}_i^i + (1/\delta)(v_i(t) - \tilde{v}_i^i) - (1/\delta - 1)(1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \geq \mu_i \\
& \Leftrightarrow (1/\delta) \left[ v_i(t) - \tilde{v}_i^i - (1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \right] \\
& \quad \geq \mu_i - \tilde{v}_i^i - (1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \\
& \Leftrightarrow \delta \geq \frac{\tilde{v}_i^i - v_i(t) + (1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i)}{\tilde{v}_i^i - \mu_i + (1/\lambda_i) \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i)} \\
& \Leftrightarrow \delta \geq \frac{1}{1 + \lambda_i(v_i(t) - \mu_i) / \left[ \lambda_i(\tilde{v}_i^i - v_i(t)) + \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \right]} \\
& \Leftrightarrow \delta_\mu \geq \frac{1}{1 + \lambda_i(v_i(t) - \mu_i) / \left[ \lambda_i(\tilde{v}_i^i - v_i(t)) + \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(b|\tilde{\mathbf{a}}^i) \right]} \\
& \Leftrightarrow \delta_\mu \geq \frac{1}{1 + d_i(v(t))}.
\end{aligned}$$

According to the first half of the proof about necessity, the above  $\delta_\mu$  in Condition (iv) is calculated by solving the optimization problem (11), which is equivalent to

$$\delta_\mu = \max_{v \in V_\mu} \min_{j \in N} \frac{1}{1 + d_j(v)}.$$

From the above, we have  $\delta_\mu \geq \min_{j \in N} \frac{1}{1 + d_j(v)}$  for any  $v \in V_\mu$ . Under the given  $v$ , the active player  $i$  is chosen such that  $d_i(v)$  is the largest (i.e.  $\frac{1}{1 + d_i(v)}$  is the smallest). Hence, we have  $\delta_\mu \geq \min_{j \in N} \frac{1}{1 + d_j(v)} = \frac{1}{1 + d_i(v)}$ . This yields  $\gamma_i(g) \geq \mu_i$ .  $\square$

**Proof of Theorem 3** Propositions 1, 2 show that Conditions (i), (ii) in Theorem 2 are necessary conditions for the existence of an efficient PPE for *any* discount factor. Suppose therefore that Conditions (i), (ii) are satisfied. It is easily checked that the following definitions of  $\mu_1^*, \mu_2^*$  guarantee that Condition (iii) of Theorem 2 are satisfied:

$$\mu_1^* = \tilde{v}_1^2 + \alpha(2, 1) [1 - \rho(b|\tilde{\mathbf{a}}^2)], \quad \mu_2^* = \tilde{v}_2^1 + \alpha(1, 2) [1 - \rho(b|\tilde{\mathbf{a}}^1)].$$

Finally, if

$$\delta \geq \delta^* \triangleq \left( 1 + \frac{1 - \lambda_1 \mu_1^* - \lambda_2 \mu_2^*}{\sum_i [\lambda_i \tilde{v}_i^i + \lambda_{-i} \alpha(i, -i) \rho(b|\tilde{\mathbf{a}}^i)] - 1} \right)^{-1},$$

then Condition (iv) of Theorem 2 is also satisfied. It follows from Theorem 2 that for each  $\delta \geq \delta^*$ ,  $V_{\mu^*}$  is a self-generating set, so every target vector in  $V_{\mu^*}$  can be achieved in a PPE. Hence  $E(\delta) \supset V_{\mu^*}$  for every  $\delta \in [\delta^*, 1)$ . To see that  $V_{\mu^*} = E(\delta)$  for every  $\delta \in [\delta^*, 1)$ , simply note that for each  $\delta$  the set  $E(\delta)$  is closed and convex, hence an interval, hence of the form  $V_{\mu}$  for some  $\mu$ . However, Condition (iii) of Theorem 2 guarantees that  $\mu \geq \mu^*$  which completes the proof.  $\square$

## References

- Abreu D, Pearce D, Stacchetti E (1990) Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58(5):1041–1063
- Athey S, Bagwell K (2001) Optimal collusion with private information. *RAND Journal of Economics* 32(3):428–465
- Bharath-Kumar K, Jaffe JM (1981) A new approach to performance-oriented flow control. *IEEE Transactions on Communications* 29(4):427–435, DOI 10.1016/j.gcb.2006.11.002, URL <http://www.sciencedirect.com/science/article/pii/S0899825606001825>
- Fudenberg D, Levine DK, Maskin E (1994) The folk theorem with imperfect public information. *Econometrica* 62(5):997–1039
- Fudenberg D, Levine DK, Takahashi S (2007) Perfect public equilibrium when players are patient. *Games and Economic Behavior* 61(1):27 – 49
- Green EJ, Porter RH (1984) Noncooperative collusion under imperfect price information. *Econometrica* 52(1):87–100
- Kandori M (1992) The use of information in repeated games with imperfect monitoring. *Review of Economic Studies* 59:581–593
- Mailath G, Samuelson L (2006) *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, Oxford, U.K.
- Mailath G, Obara I, Sekiguchi T (2002) The maximum efficient equilibrium payoff in the repeated prisoners' dilemma. *Games and Economic Behavior* 40(1):99–122
- Radner R, Myerson R, Maskin E (1986) An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. *Review of Economic Studies* 53(1):59 – 69
- van der Schaar M, Xiao Y, Zame WR (2013) Designing efficient resource sharing for impatient players using limited monitoring. [Http://arxiv.org/abs/1309.0262](http://arxiv.org/abs/1309.0262)