



EIEF Working Paper 20/13
August 2013

**Designing Efficient Resource Sharing For
Impatient Players Using Limited Monitoring**

by

Mihaela van der Schaar

(Department of Electrical Engineering, UCLA)

Yuanzhang Xiao

(Department of Electrical Engineering, UCLA)

William R. Zame

(Department of Economics, UCLA)

Designing Efficient Resource Sharing For Impatient Players Using Limited Monitoring[☆]

Mihaela van der Schaar^a, Yuanzhang Xiao^b, William Zame^c

^a*Department of Electrical Engineering, UCLA. Email: mihaela@ee.ucla.edu.*

^b*Department of Electrical Engineering, UCLA. Email: yxiao@ee.ucla.edu.*

^c**Corresponding Author** *Department of Economics, UCLA, Los Angeles, CA 90095
Email: zame@econ.ucla.edu; Telephone 310-985-3091*

Abstract

The problem of efficient sharing of a resource is nearly ubiquitous. Except for pure public goods, each agent's use creates a negative externality; often the negative externality is so strong that efficient sharing is impossible in the short run. We show that, paradoxically, the impossibility of efficient sharing in the short run enhances the possibility of efficient sharing in the long run, even if outcomes depend stochastically on actions, monitoring is limited and users are not patient. We base our analysis on the familiar framework of repeated games with imperfect public monitoring, but we extend the framework to view the monitoring structure as chosen by a designer who balances the benefits and costs of more accurate observations and reports. Our conclusions are much stronger than in the usual folk theorems: we do not require a rich signal structure or patient users and provide an explicit online construction of equilibrium strategies.

Keywords: repeated games, imperfect public monitoring, perfect public equilibrium, efficient outcomes, resource allocation games

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.

1. Introduction

The problem of efficient sharing of a resource – a physical resource, a prize, a market – is nearly ubiquitous. Unless the resource is a pure public good, each agent’s use of the resource imposes a negative externality on other users. Hence (self-interested, strategic) agents will find it difficult to share the resource efficiently, at least in the short run. In some circumstances – those we focus on in this paper – the negative externality is so strong – competition for the resource is so destructive – that it will be *impossible* for users to share the resource efficiently, at least in the short run. The purpose of this paper is to show that – perhaps paradoxically – the impossibility of efficient sharing in the short run enhances the possibility of efficient sharing in the long run – even when outcomes depend stochastically on actions, monitoring is very limited and players are not very patient.

We formalize our analysis using the familiar framework of repeated games with imperfect public monitoring but with important differences in both the formulation and the conclusions. With respect to the formulation, the important difference is in the way we view the monitoring structure. In the usual models of games with imperfect public monitoring, the monitoring structure is viewed as exogenous and fixed. In the canonical model of Green and Porter (1984) for instance, the players compete in a Cournot quantity-setting game but receive feedback only about market prices (rather than quantity choices of other firms) which are determined by random market demand. We are motivated by the many situations in which the feedback received by the players arises from the action choices of a strategic actor – the *designer* – who must weigh (among other considerations) the trade-offs between more

accurate observations of player actions and more accurate reports provided to players about those observations on the one hand and the costs and consequences of observations and reports on the other hand. Consider for instance a repeated contest (for details see Example 2 in Section 3). In each period, players choose effort levels which determine (stochastically) the contest winner. The designer (in this case the contest operator) does not observe effort – and so certainly cannot announce it – but does observe the identity of the winner, and could announce that. However announcing the identify of the winner would violate the privacy of the winner and the losers; if privacy is valued, the designer must weigh the trade-off between the value of maintaining privacy and the (possible) efficiency gain of making more information public. Because we wish to emphasize the role played by this and similar *choices* of the monitoring structure we formalize an elaborated model in which the choice of the monitoring structure – both what the designer observes and what the designer announces – is made explicit. However, the reduced form that results from our elaborated model once the designer has chosen a monitoring structure looks just the same as the reduced form that is familiar from the standard model of repeated games with imperfect public monitoring.

With respect to conclusions, we cite three important differences: we do not assume a rich signal structure (rather, we require only two signals), we do not assume players are arbitrarily patient (rather, we find an explicit lower bound on the requisite discount factor), and we provide an *explicit (distributed) algorithm* that takes as inputs the parameters – stage game payoffs, discount factor, target payoff – and computes the strategy – the action to be

chosen by each player following each public history. This algorithm can be carried out by each player separately and in real time – there is no need for the designer to specify/describe the strategies to be played. A consequence of our constructive algorithm is that the strategies we identify enjoy a useful robustness property: generically, the equilibrium strategies are, for many periods, locally constant in the parameters of the environment and of the problem.

Within our structure, we abstract what we see as the essential features of the resource allocation problems 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 allocation scenario, $\tilde{\mathbf{a}}^i$ would be the profile in which only player i accesses the resource.) 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 $\{U(\tilde{\mathbf{a}}^i)\}$. (In the resource allocation scenario, this corresponds to the assumption that allowing access to the resource by more than one individual strictly lowers (weighted) social welfare.) We capture the notion that monitoring is very limited by assuming 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} , and that (profitable) single-player deviations from i 's preferred action profile $\tilde{\mathbf{a}}^i$ can be statistically distinguished from conformity with $\tilde{\mathbf{a}}^i$ in the same way. (But we do not assume that different deviations from $\tilde{\mathbf{a}}^i$ can be distinguished from *from each other*. For further comments, see Examples 2 and 3 in Section 3.) We emphasize the setting in which there are only two signals – “good” and “bad” – because

this setting offers the sharpest results and the clearest intuition and, as we shall see, because two signals are often enough. To help understand the commonplace nature of our problem and assumptions, we offer three examples: the first is a repeated prisoner’s dilemma (although with lower cooperative payoffs than usual), the second is a repeated contest, the third is a repeated resource sharing game.

Not surprisingly we build on the framework of (Abreu, Pearce, and Stacchetti (1990); hereafter APS). Our main technical result (Theorem 1) provides conditions (on the information and payoff structures and the discount factor) that are both *necessary and sufficient* for the set of payoffs that guarantee each player a given level of security to be self-generating. Because every payoff vector in a self-generating set can be supported in a perfect public equilibrium (PPE), this leads immediately to sufficient conditions for the same sets to consist of payoff vectors that can be achieved in PPE, and an algorithm for the corresponding PPE strategies (Theorem 2). Our robustness conclusion (Theorem 3) follows from the nature of the algorithm. For games with two players, other considerations lead to the conclusion that maximal sets of PPE payoffs must have a special form and so thus to a characterization of the maximal set of PPE payoffs (Theorem 4). A surprising aspect of this characterization is that there is a discount factor $\delta^* < 1$ such that any efficient payoff that can be achieved as a PPE payoff for *some* discount factor δ can already be achieved as a PPE payoff as soon as the discount factor δ exceeds *some threshold* δ^* . Patience is rewarded – but only up to a point.¹

¹Mailath, Obara, and Sekiguchi (2002) establish a similar result for the repeated Prisoner’s Dilemma with perfect monitoring; Athey and Bagwell (2001) establish a parallel

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. The first and most obvious comparison is with (Fudenberg, Levine, and Maskin (1994); hereafter FLM) on the Folk Theorem for repeated games with imperfect public monitoring. As do we, FLM consider a situation in which a single stage game G with action space \mathbf{A} and utility function $U : \mathbf{A} \rightarrow \mathbb{R}^n$ is played repeatedly over an infinite horizon; monitoring is public but imperfect, so players do not observe actions but only a public signal of those actions. In this setting, $\text{co}[U(\mathbf{A})]$ is the closure of the set of payoff profiles that can be achieved as long run average utilities for *some* discount factor and *some* infinite set of plays of the stage game G . Under certain assumptions, FLM prove that any payoff vector in the interior of $\text{co}[U(\mathbf{A})]$ that is strictly individually rational can be achieved in a PPE of the infinitely repeated game. However, the assumptions FLM maintain are very different from ours in two very important dimensions (and some other dimensions that seem less important, at least for the present discussion). The first is that the signal structure is rich and informative; in particular, that the number of signals is at least one less than the number of actions of any two players. The second is that players are arbitrarily patient: that is, the discount factor δ is as close to 1 as we like. (More precisely: given a target utility profile v , there is some $\delta(v)$ such that if the discount factor $\delta > \delta(v)$ then there is a PPE of the

result for symmetric equilibrium payoffs of two-player symmetric repeated Bertrand games. We are unaware of any general results that have this flavor.

repeated game that yields the target utility profile v .) In particular, FLM do not identify any PPE for any *given* discount factor $\delta < 1$. By contrast, we require only two signals *even if action spaces are infinite* and we do *not* assume players are patient: all target payoffs can be achieved for some *fixed* discount factor – which may be very far from 1. Moreover, because FLM consider only payoffs in the interior of $\text{co}[U(\mathbf{A})]$, they have nothing to say about achieving *efficient* payoffs. Their results do imply that efficient payoffs can be arbitrarily well approximated by payoffs that can be achieved in PPE, but only if the corresponding discount factors are arbitrarily close to 1. By contrast, (Fudenberg, Levine, and Takahashi (2007); hereafter FLT) *do* show how (some) efficient payoffs can be achieved in PPE. Given Pareto weights $\lambda_1, \dots, \lambda_n$ set $\Lambda = \sup\{\sum \lambda_i U_i(\mathbf{a}) : \mathbf{a} \in \mathbf{A}\}$ and consider the hyperplane $H = \{x \in \mathbb{R}^n : \sum \lambda_i x_i = \Lambda\}$. The intersection $H \cap \text{co}[U(\mathbf{A})]$ is a part of the Pareto boundary of $\text{co}[U(\mathbf{A})]$. As do we, FLT ask what vectors in $H \cap \text{co}[U(\mathbf{A})]$ can be achieved in PPE of the infinitely repeated game. They identify the largest (compact convex) set $Q \subset H \cap \text{co}[U(\mathbf{A})]$ 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, because FLT consider arbitrary stage games and arbitrary monitoring structures, the set Q identified by FLT may be empty, and FLT do not provide any conditions that guarantee that Q is not empty. Moreover, as in FLM, FLT assume that players are arbitrarily patient, so do not identify any PPE for any *given* discount factor $\delta < 1$. Having said this, we should also point out that FLT identify the closure of the set of *all* payoff vectors in the interior of $H \cap \text{co}[U(\mathbf{A})]$ that can be achieved

in a PPE for some discount factor, while we identify only some. So there is a trade-off: FLT find more PPE payoffs but provide much less information about the ones they find; we find fewer PPE payoffs but provide much more information about the ones we find.

At the risk of repetition, we want to emphasize the most important features of our results. The first is that we do not assume discount factors are arbitrarily close to 1. 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 impose different – and in many ways weaker – requirements on the monitoring structure; indeed, we require only two signals, even if 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 by a designer. In the latter case it seems obvious – and in practice may be of supreme importance – that the designer 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 (and preserving privacy as well). More generally, the designer may face a trade-off between the efficiency obtainable with a finer information structure and the cost of using that information structure. (We will return to this point later.) Finally, because we provide a distributed algorithm for calculating equilibrium play, neither the agents nor a designer need to work out the equilibrium strategies in advance; all calculations can be done online, in real time.

Following this Introduction, Section 2 presents the formal model; Section

3 presents three examples that illustrate the model. Section 4 presents some preliminary results, presenting conditions under which *no* efficient payoffs can be achieved in PPE for *any* discount factor. Section 5 presents the main technical result (Theorem 1); Section 6 presents the implications for PPE (Theorems 2,3) and a comparison with FLT; Section 7 specializes to the case of two players (Theorem 4). Section 8 returns to the examples to illustrate both the conclusions and the general framework. Section 9 concludes. We relegate all proofs to the Appendix.

2. Model

The reduced form of our model will closely resemble the familiar framework of a repeated game with imperfect public monitoring and we state and prove our formal results in the context of that reduced form. However, because we want to emphasize the role played by the designer, we begin by presenting a more elaborated form.

2.1. Stage Game: Elaborated Form

There are $n + 1$ (potential) actors in our framework: n *players* and a *designer*. Players are characterized by an (exogenously given) *game form*:

- a (measurable) space Z of *outcomes*
- for each player i
 - a (measurable) space A_i of *actions*
 - a (measurable) *utility function* $u_i : A_i \times Z \rightarrow \mathbb{R}$
- a (measurable) mapping $\mathbf{a} \mapsto \pi(\cdot|\mathbf{a}) : \mathbf{A} = A_1 \times \cdots \times A_n \rightarrow \Delta(Z)$

We view $\pi(z|\mathbf{a})$ as the probability that the outcome $z \in Z$ occurs when players choose the action profile $\mathbf{a} \in \mathbf{A}$. Thus the joint actions of players $\mathbf{a} \in \mathbf{A}$ stochastically determine an outcome $z \in Z$, and each player's realized utility depends on its own action and the realized outcome.² For the moment we require only that the spaces A_i , the utility functions u_i and the probability mapping π be measurable, so that utilities in the reduced form be defined; but later we will insist that the spaces be compact metric and that the utility functions and the probability mapping be continuous.

The designer is characterized by a *monitoring technology*:

- a set of Φ of pairs (X, φ) where:
 - X is a (measurable) space
 - $z \mapsto \varphi(\cdot|z) : Z \rightarrow \Delta(X)$ is a (measurable) mapping

A pair (X, φ) is a *measurement device*.

- a set Ψ of pairs (Y, ψ) where
 - Y is a (measurable) space
 - $x \mapsto \psi(\cdot|x) : X \rightarrow \Delta(Y)$ is a (measurable) mapping

A pair (Y, ψ) is an *announcement rule*.

For the moment, we again require only that the spaces X, Y and the mappings φ, ψ be measurable, but later we will insist that the spaces be compact metric and that the mappings be continuous. Given a choice $(X, \varphi) \in \Phi$ we interpret

²We could incorporate actions into the space Z of outcomes so that realized utility depended only on outcomes, but it seems useful to keep separate track of own actions.

$\varphi(x|z)$ as the probability that the designer measures (observes) x when the outcome z has actually occurred. Given a choice $(Y, \psi) \in \Psi$, we interpret $\psi(y|x)$ as the probability that the designer makes the (*public*) *announcement* of the signal $y \in Y$ when the observation x has actually been made. A pair of choices $(X, \varphi) \in \Phi, (Y, \psi) \in \Psi$ constitute the *monitoring structure*.

2.2. Stage Game: Reduced Form

The *reduced form* of the stage game consists of

- a set $N = \{1, \dots, n\}$ of players
- for each player i
 - a (measurable) space A_i of actions
 - a (measurable) utility function $U_i : \mathbf{A} = A_1 \times \dots \times A_n \rightarrow \mathbb{R}$
- a (measurable) compact metric space of public signals Y
- a (measurable) map $\mathbf{a} \mapsto \rho(\cdot|\mathbf{a}) : \mathbf{A} \rightarrow \Delta(Y)$

We interpret $U_i(\mathbf{a})$ as i 's *ex ante* (expected) utility when \mathbf{a} is played and $\rho(y|\mathbf{a})$ as the probability that the signal y is observed when \mathbf{a} is played.

2.3. Stage Game: From the Elaborated Form to the Reduced Form

To pass from the elaborated form to the reduced form we simply define the *ex ante* (expected) utilities $U_i(\mathbf{a})$ and the probability distribution $\rho(\cdot|\mathbf{a})$ over public signals as functions of the action profile \mathbf{a} that is played.

For $\mathbf{a} \in \mathbf{A}$ and $D \subset Y$ these are:

$$U_i(\mathbf{a}) = \int_Z u_i(a_i, z) d\pi(z|\mathbf{a})$$

$$\rho(D|\mathbf{a}) = \int_Y \int_X \int_Z \mathbf{1}_D d\psi(y|x) d\varphi(x|z) d\pi(z|\mathbf{a})$$

If Z, X, Y are all finite the last equation can be re-written more simply as

$$\rho(y|\mathbf{a}) = \sum_{x \in X} \sum_{z \in Z} \psi(y|x) \varphi(x|z) \pi(z|\mathbf{a})$$

Under the maintained assumptions on realized utility, outcome mapping, measurement technology and announcement rules, the derived *ex ante* utilities and signal distribution are measurable; if the former are continuous, so are the latter.

2.4. The Repeated Game with Imperfect Public Monitoring

In the repeated game, the reduced stage game G is played in every period $t = 0, 1, 2, \dots$. Given the signal structure, 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, the actions taken by player i , and the realized utilities observed by player i , so a *private history* of length t is a sequence $(a_i^0, \dots, a_i^{t-1}; u_i^0, \dots, u_i^{t-1}; y^0, \dots, y^{t-1}) \in A_i^t \times \mathbb{R}^t \times Y^t$. 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 the set of pure 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/utility history; equivalently, a mapping from public histories to i 's pure actions $\sigma_i : \mathcal{H} \rightarrow A_i$.

We assume all players discount future utilities using the same discount factor $\delta \in (0, 1)$ and we use long-run averages, so if the stream of expected utilities is $\{u^t\}$ the vector of long-run average utilities is $(1 - \delta) \sum_{t=0}^{\infty} \delta^t u^t$. 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 σ .

As usual a strategy profile σ 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.

2.5. Interpretation

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 Mailath and Samuelson (2006), this assumption admits a number of possible interpretations, each of which is appropriate in

some circumstances. The first is that utility is not realized until the game terminates. The second is that the outcome z and the public signal y coincide, so that realized utility depends only on own action and the public signal (both of which are observed). The third is that – at least in the equilibria and deviations under consideration – the information provided by realized utility is already provided by the public signal. (See Example 2 below.) A fourth is that even if utility *is* realized during play *and* realized utility *does* provide information not provided by the public signal, this additional information is not used. Lest this last interpretation seems odd, recall that if players other than i follow public strategies then it is optimal for player i to follow a public strategy as well; in particular if other players make no use of information provided by their own realized utility then it is optimal for player i to make no use of information provided by i 's realized utility. (Again, see Example 2 below.) Finally, it should be kept in mind that by restricting our attention to PPE we are tying our own hands; since our objective is to support efficient sharing, restricting to a particular class of strategies only makes our results stronger.

2.6. Assumptions on the Stage Game

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.

We assume that the spaces Z, A_i, X, Y are all compact metric and that the functions/mappings u_i, π, φ, ψ are all continuous; as noted this implies that the functions/mappings U_i, ρ are continuous as well.

Set $U(\mathbf{A}) = \{U(\mathbf{a}) \in \mathbb{R}^n : \mathbf{a} \in \mathbf{A}\}$ and let $\text{co}(U(\mathbf{A}))$ be the convex hull

of $U(\mathbf{A})$. For each i set

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

Compactness of the action space \mathbf{A} and continuity of utility functions U_i guarantee that $U(\mathbf{A})$ and $\text{co}[U(\mathbf{A})]$ are compact, that \tilde{v}^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 for player i is attained at a *unique* strategy profile $\tilde{\mathbf{a}}^i$.³ 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 typically be the (unique) action profile at which agent i has optimal access to the resource and other agents have none. For this reason, we will often say that i is *active* at the profile $\tilde{\mathbf{a}}^i$ and other players are *inactive*. 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 $\text{co}(U(\mathbf{A}))$ is the closure 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 closure 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 $\text{co}[U(\mathbf{A})]$ as the set of *feasible payoffs* and to V as the set of *efficient payoffs*.⁴

³This assumption could be avoided, at the expense of some technical complication.

⁴The latter is a slight abuse of terminology: because V is the intersection of the set of feasible payoffs with a bounding hyperplane, 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.

We abstract the motivating class of resource allocation problems by imposing conditions on the set of preferred utility profiles. The first is made largely for convenience (and is generically satisfied whenever action spaces are finite); the second abstracts the idea that there are strong negative externalities.

Assumption 1 The vectors $\tilde{v}^1, \dots, \tilde{v}^n$ are linearly independent.

Assumption 2 The affine span of \tilde{V} is a hyperplane H and all *ex ante* utility vectors of the game other than the those in \tilde{V} lie below H . That is, there are weights $\lambda_1, \dots, \lambda_n > 0$ such that $\sum \lambda_j u_j(\tilde{\mathbf{a}}^i) = 1$ for each i and $\sum \lambda_j u_j(\mathbf{a}) < 1$ for each $\mathbf{a} \in \mathbf{A}, \mathbf{a} \notin \tilde{\mathbf{A}}$.⁵

2.7. Assumptions on the Monitoring Structure

As noted in the Introduction, we focus on the case in which there are only two signals.

Assumption 3 The set Y contains precisely two signals and $\rho(y|\mathbf{a}) > 0$ for every $y \in Y$ and $\mathbf{a} \in \mathbf{A}$. (The monitoring structure has *full support*.)

We assume that profitable deviations from the profiles $\tilde{\mathbf{a}}^i$ exist and be statistically detected in a particularly simple way.

Assumption 4 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, there is a labeling $Y = \{y_g^i, y_b^i\}$

⁵That the sum is 1 is just a normalization.

with the property that

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

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 y_g^i is observed (equivalently: strictly increases the probability that the “bad” signal y_b^i is observed).

The import of Assumption 4 is 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 – and for the possibility that which signal is “good” and which is “bad” depend on the identity of the *active* player i .

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. We present the first example directly in the reduced form and the other two examples in both the elaborated and reduced forms.

Example 1: A Repeated Prisoners’ Dilemma

We begin by discussing a simple Prisoner’s Dilemma but with a payoff structure slightly different from the familiar one; see Table 1. For our pur-

poses we assume $B > 2c > 2b > 0$. As usual, (D, D) is a strictly dominant strategy profile; the difference between the payoffs shown here and the usual ones is that (C, C) is Pareto dominated by randomizing between (C, D) and (D, C) . See Figure 1.

There are two signals: $Y = \{y_g, y_b\}$; the probability distribution over signals following actions is

$$\pi(y_g|\mathbf{a}) = \begin{cases} p & \text{if } \mathbf{a} = (C, C) \\ q & \text{if } \mathbf{a} = (C, D) \text{ or } (D, C) \\ r & \text{if } \mathbf{a} = (D, D) \end{cases} \quad (1)$$

where $p, q, r \in (0, 1)$; for our purposes we assume $p \geq q > r$. It is easily checked that the stage game and monitoring structure satisfy our assumptions. (Note that y_g is the good signal for both players.) As we will show in Section 4, we can completely characterize the most efficient outcomes that can be achieved in a PPE. To summarize the conclusion, for each discount factor $\delta \in (0, 1)$ write $E(\delta)$ for the set of efficient (average) payoffs that can be achieved when the discount factor is δ . Set

$$\delta^* = \frac{1}{1 + \left(\frac{B - 2\frac{q}{q-r}b}{B + 2\frac{1-q}{q-r}b} \right)}$$

It follows from Theorem 4 that if $\delta \geq \delta^*$ then

$$E(\delta) = \{(v_1, v_2) : v_1 + v_2 = B; v_i \geq q/(q-r)b\}$$

Note that the set of efficient equilibrium outcomes *does not* increase as $\delta \rightarrow 1$;

Table 1: Modified Prisoners' Dilemma

	C	D
C	(c, c)	$(0, B)$
D	$(B, 0)$	(b, b)

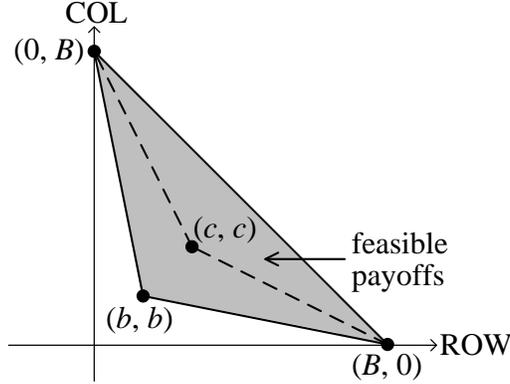


Figure 1: Feasible Region for the Modified Prisoners' Dilemma

as we noted in the Introduction, patience is rewarded but only up to a point. See Figure 5.

Example 2: A Repeated Contest

We consider a repeated contest. In each period, a set of $n \geq 2$ players competes for the use of a single indivisible resource/prize each of them values at $R > 0$. Winning the contest depends (stochastically) on the effort exerted by each player; we write $A_i = [0, 1]$ for the set of i 's effort levels (actions). 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. If $\mathbf{a} = (a_i)$ is the vector of effort levels then the probability agent i obtains the wins the contest (obtains the resource/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. The assumption that $\eta < 1$ reflects that there is always some probability the prize is not awarded; κ 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 the prize. Utility is separable in reward and effort; effort is costly with constant marginal cost $c > 0$. To avoid trivialities and conform with Assumptions 1-4 we assume $R\eta > c$ and that $\kappa > \frac{1}{2} \left(\eta - \frac{c}{R} \right)$.

In the elaborated form of the stage game, players are $N = \{1, \dots, n\}$, action sets are $A_i = [0, 1]$, outcomes are $Z = \{z_0, \dots, z_n\}$ (where z_0 is interpreted as “no one wins” and z_i is interpreted as “ i wins”) and i 's realized utility as a function of his own effort level a_i and the outcome z is

$$u_i(a_i, z_k) = \begin{cases} R - ca_i & \text{if } k = i \\ -ca_i & \text{if } k \neq i \end{cases}$$

In this context it seems natural to assume that the designer observes who wins – how else could the prize be awarded? – so that $X = Z$ and φ is the identity. We assume that the designer wishes to preserve privacy so announces only whether or not *some* player won the contest but not *the identity of the winner*. Hence the reporting rule $(Y^1, \psi^1), (Y^2, \psi^2)$, where

- $Y^1 = \{y_b, y_g\}$; $\psi^1(z_k) = y_b$ if $k = 0$, $\psi^1(z_k) = y_g$ if $k \neq 0$
- $Y^2 = Z$; $\psi^2(z_k) = z_k$ for all $k = 0, \dots, n$

In the first case, the designer announces whether or not there has been a winner; in the second case the designer also announces the identity of the winner.

In the reduced forms of the stage game, the *ex ante* expected utilities are given by

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

In the first case, the signal distribution is

$$\bar{\rho}(y_*|a) \begin{cases} 1 - \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } * = b \\ \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } * = g \end{cases}$$

In the second case the signal distribution is

$$\hat{\rho}(y_k|a) = \begin{cases} 1 - \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } k = 0 \\ a_k \left(\eta - \kappa \sum_{j \neq k} a_j \right)^+ & \text{if } k \neq 0 \end{cases}$$

Straightforward but somewhat messy calculations show that in either case the reduced form satisfies all of our assumptions. (Player i 's preferred action profile $\tilde{\mathbf{a}}^i$ has $\tilde{\mathbf{a}}_i^i = 1$ and $\tilde{\mathbf{a}}_j^i = 0$ for $j \neq i$: i exerts maximum effort, others exert none. Note that this does not guarantee that i wins the contest – there may still be no winner – but the effort profiles $\tilde{\mathbf{a}}^i$ are precisely those that maximize the probability that *someone* wins the prize.)

The first reporting rule preserves privacy, the second rule does not. How-

ever, the second reporting rule provides more information to players. Suppose for instance that a strategy profile σ calls for $\tilde{\mathbf{a}}^i$ to be played after a particular history. If all players follow σ 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. Which reporting rule – hence which monitoring structure – should be chosen by the designer will depend on the tradeoff the designer makes between preserving privacy and sustaining a wider range of PPE. We will see a similar but even starker tradeoff in Example 3 following.

Example 3: Resource Sharing

We consider $n \geq 3$ users (players) who 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 ε is distributed in some interval $[0, \bar{\varepsilon}]$ with (known) distribution ν . 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 ε 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", namely the ratio of the p -th power of its own packet rate to the delay:

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

where $p > 0$ is a parameter that represents trade-off between rate and delay.⁶ (If delay is infinite utility is 0.) Formally, we identify the outcome with the pair consisting of the vector \mathbf{a} of packet rates and the realized shock ε , so $Z = A \times [0, \bar{\varepsilon}]$ and $\pi(\cdot|\mathbf{a}) = \delta_{\mathbf{a}} \times \nu$ where $\delta_{\mathbf{a}}$ is point mass at \mathbf{a} and ν is the given distribution of shocks.

The designer does not observe packet rates but can measure the delay, but with error and at a cost. Thus the space of measurements is $X = [0, \infty]$ and the measurement technology consists of a space of maps $(\mathbf{a}, \varepsilon) \mapsto \varphi(\cdot|(\mathbf{a}, \varepsilon)) : \mathbf{A} \times [0, \bar{\varepsilon}] \rightarrow \Delta(X)$. Many possible reporting technologies are possible; we assume the designer reports only whether the measured delay was above or

⁶In order to guarantee that the reduced form satisfies our assumptions we assume $\bar{\varepsilon} \leq \frac{2}{2+p}\chi$.

below a chosen threshold d_0 ; say $Y = \{y_\ell, y_h\}$ where y_ℓ is interpreted as “delay was low (below d_0)” and y_h is interpreted as “delay was high (above d_0).”

In the reduced form, each player i 's ex-ante payoff is

$$U_i(\mathbf{a}) = \begin{cases} a_i^p (\chi - \frac{\bar{\varepsilon}}{2} - \sum_{j=1}^n a_j) & \text{if } \sum_{j=1}^n a_j \leq \chi - \bar{\varepsilon} \\ a_i^p (\chi - \sum_{j=1}^n a_j) \frac{\chi - \sum_{j=1}^n a_j}{2\bar{\varepsilon}} & \text{if } \chi - \bar{\varepsilon} < \sum_{j=1}^n a_j < \chi \\ 0 & \text{otherwise} \end{cases}$$

and the distribution of signals is

$$\rho(y_\ell | \mathbf{a}) = \int_0^{\chi - \sum_{j=1}^n a_j - \frac{1}{d_0}} d\nu(x) = \frac{[\chi - \sum_{j=1}^n a_j - \frac{1}{d_0}]_0^{\bar{\varepsilon}}}{\bar{\varepsilon}},$$

where $[x]_a^b \triangleq \min\{\max\{x, a\}, b\}$ is the projection of x in the interval $[a, b]$. Note that y_ℓ 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.

It might seem to the reader that the players could back out realized delay from their own realized utility and hence that announcements are irrelevant – but this is not quite so. Players who choose packet rates greater than 0 *can* back out realized delay from their own realized utility but at any one of the preferred action profiles $\tilde{\mathbf{a}}^i$ and at any single-player deviation from any one of the preferred action profiles $\tilde{\mathbf{a}}^i$, at least one player will choose a packet rate $a_j = 0$ and hence will experience realized utility $U_i(\mathbf{a}) = 0$; that player *cannot* back out observed delay. Hence announcements serve

to (statistically) inform players *who have complied* of the existence of some player *who has not complied*. Put differently, announcements serve to keep all players on the same informational page.

4. Ruling out Some Efficient PPE Payoffs

Throughout this Section, we consider a fixed reduced form and maintain the notation and assumptions of Section 2. Our ultimate goal is to find conditions – on the discount factor among other things – that enable us to construct PPE that achieve payoffs in V (efficient payoffs).

We first show that under certain conditions, certain efficient payoffs *cannot* be achieved in PPE no matter what the discount factor is. To this end, we identify two measures of benefits from deviation. (These same measures will play a prominent role in the next Section as well.) Given $i, j \in N$ with $i \neq j$ set:

$$\alpha(i, j) = \sup \left\{ \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(y_b^i | a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)} : \right. \\ \left. a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i) \right\} \quad (2)$$

$$\beta(i, j) = \inf \left\{ \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(y_b^i | a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)} : \right. \\ \left. a_j \in A_j, u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) < u_j(\tilde{\mathbf{a}}^i), \rho(y_b^i | a_j, \tilde{\mathbf{a}}_{-j}^i) < \rho(y_b^i | \tilde{\mathbf{a}}^i) \right\} \quad (3)$$

(We follow the usual convention that the supremum of the empty set is $-\infty$ and the infimum of the empty set is $+\infty$.)

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 $\rho(y_b^i|a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i|\tilde{\mathbf{a}}^i)$ 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 strictly increase the probability that the bad signal occurs, so $\alpha(i, j)$ is either $-\infty$ or strictly positive. In the definition of $\beta(i, j)$ we consider only deviations that are strictly unprofitable *and* strictly decrease the probability that the bad signal occurs, so $\beta(i, j)$ is the infimum of strictly positive numbers and so is necessarily $+\infty$ or finite and non-negative.⁷

To understand the significance of these numbers, think about how player j could gain by deviating from $\tilde{\mathbf{a}}^i$. Most obviously, 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. However, 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.

Because \tilde{V} lies in the supporting hyperplane H and the utilities for ac-

⁷Note that if we strengthened Assumption 4 so that *any* deviation – profitable or not – increased the probability of a bad signal (as is the case in Examples 1-3 and would be the case in most resource allocation scenarios), then $\beta(i, j)$ would be the infimum of the empty set whence $\beta(i, j) = +\infty$.

tion profiles not in $\tilde{\mathbf{A}}$ lie strictly below H , in order that the strategy profile σ achieves an efficient payoff it is necessary and sufficient that σ use only preferred action profiles: $U(\sigma) \in V$ if and only if $\sigma(h) \in \tilde{\mathbf{A}}$ for every public history h (independently of the discount factor δ). For PPE strategies we can say a lot more. The first Proposition is almost obvious; the second and third seem far from obvious. (All proofs are in the Appendix.)

Proposition 1. *In order that \tilde{v}^i be achievable in a PPE equilibrium (for any discount factor δ) it is necessary and sufficient that $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) \leq u_j(\tilde{\mathbf{a}}^i)$ for every $j \neq i$ and every $a_j \in A_j$.*

Proposition 2. *If σ is an efficient PPE (for any discount factor δ) and i is active following some history (i.e., $\sigma(h) = \tilde{\mathbf{a}}^i$ for some h) then*

$$\alpha(i, j) \leq \beta(i, j) \tag{4}$$

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

Proposition 3. *If σ is an efficient PPE (for any discount factor δ) and i is active following some history (i.e., $\sigma(h) = \tilde{\mathbf{a}}^i$ for some h) then*

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(y_b^i | a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)] \tag{5}$$

The import of Propositions 2 and 3 is that if any of these inequalities fail then certain efficient payoff vectors can *never* be achieved in PPE, no

matter what the discount factor is. In the next Sections, we show how these inequalities and other conditions yield necessary and sufficient conditions that certain sets be self-generating and hence yield sufficient conditions for efficient PPE.

Proposition 2 might seem quite mysterious: α is a measure of the current gain to deviation and β is a measure of the future gain to deviation; there seems no obvious reason why PPE should necessitate any particular relationship between α and β . As the proof will show, however, the assumption of two signals and the efficiency of payoffs in V imply that α is bounded above and β is bounded below by the same quantity, which is a weighted difference of continuation values – a quantity that does have an obvious connection to PPE.

5. Characterizing Efficient Self-Generating Sets

As in the previous Section, we consider a fixed reduced form and maintain the notation and assumptions of Section 2. In order to find efficient PPE payoffs we follow APS and look for self-generating sets of efficient payoffs.

Fix a subset $W \subset \text{co}[U(\mathbf{A})]$ and a *target payoff* $v \in \text{co}[U(\mathbf{A})]$. Recall from APS that v can be *decomposed with respect to* W (for a given 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 $a_j \in A_j$

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

Write $\mathcal{B}(W, \delta)$ for the set of target payoffs $v \in \text{co}[U(\mathbf{A})]$ that can be decomposed with respect to W (for the discount factor δ). Recall that 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 .

Because V lies in the hyperplane H , if $v \in V$ and it is possible to decompose $v \in V$ with respect to *any* set and for *any* discount factor, then the associated action profile \mathbf{a} must lie in $\tilde{\mathbf{A}}$ and the continuation payoffs must lie in V . 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 $\tilde{\mathbf{a}}^i$. For player $j \neq i$, a profitable deviation 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 without further comment that $V_\mu \neq \emptyset$. For lack of a better term, we say that V_μ is *regular* if for each $i \in N$ there is a vector $\hat{v}^i \in V_\mu$ such that $\hat{v}_j^i = \mu_j$ for each $j \neq i$. Whether or not V_μ is regular depends both on the shape of V and on the magnitude of μ : see Figures 2, 3, 4 for instance. A few simple facts are useful to note:

- If $\tilde{v}_j^i = 0$ for all $i, j \in N$ with $i \neq j$ (as is the case in many resource sharing scenarios such as Examples 2, 3) then V_μ is regular for every $\mu \geq 0$.
- If $V_\mu \neq \emptyset$ and V_μ is a subset of the interior of V (relative to the hyperplane H) then V_μ is regular.
- If v lies in the interior of V (relative to the hyperplane H) and $\mu = v - \epsilon \cdot \mathbf{1}$ for $\epsilon > 0$ sufficiently small, then $v \in V_\mu$ and V_μ is regular.
- If V_μ is not a singleton then it must contain a point of the interior of V (relative to the hyperplane H).

If V_μ is a singleton, it can only be a self-generating set (and hence achievable in a PPE) if $V_\mu = \tilde{v}^i$ for i ; because we have already characterized this possibility in Proposition 1, we focus on the non-degenerate case in which V_μ is not a singleton and hence contains a point of the interior of V . Note that a

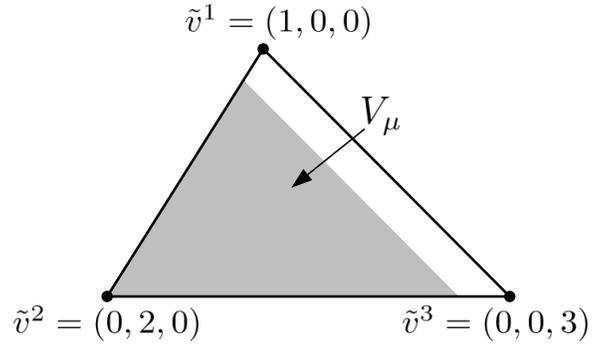


Figure 2: $\mu = (0, 1/4, 0)$; V_μ is regular

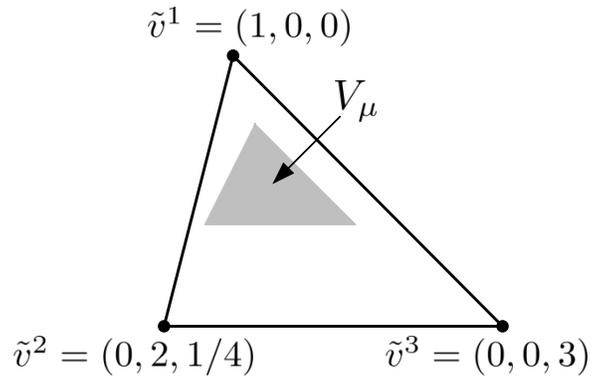


Figure 3: $\mu = (1/2, 1/2, 1/2)$; V_μ is regular

point in the interior of V can only be achieved by a repeated game strategy in which *all* players are active following some history.

The following result provides necessary and sufficient conditions on μ , the payoff structure, the information structure and the discount factor that a regular V_μ be a self-generating set.

Theorem 1. *Fix μ ; assume that V_μ is regular and not an extreme point of V . In order that V_μ be a self-generating set, it is necessary and sufficient that the following conditions be satisfied:*

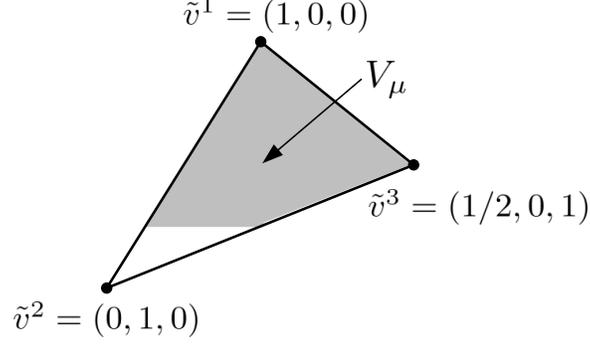


Figure 4: $\mu = (1/4, 0, 0)$; V_μ is not regular

Condition 1 for all $i, j \in N$ with $i \neq j$:

$$\alpha(i, j) \leq \beta(i, j) \quad (6)$$

Condition 2 for all $i \in N$ and all $a_i \in A_i$:

$$\tilde{v}_i^i - u_i(a_i, \tilde{\mathbf{a}}_{-i}^i) \geq \frac{1}{\lambda_i} \sum_{j \neq i} \lambda_j \alpha(i, j) [\rho(y_b^i | a_i, \tilde{\mathbf{a}}_{-i}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)] \quad (7)$$

Condition 3 for all $i \in N$:

$$\mu_i \geq \max_{j \neq i} (\tilde{v}_i^j + \alpha(j, i)[1 - \rho(y_b^j | \tilde{\mathbf{a}}^j)]) \quad (8)$$

Condition 4 the discount factor δ satisfies:

$$\delta \geq \underline{\delta}_\mu \triangleq \left(1 + \frac{1 - \sum_i \lambda_i \mu_i}{\sum_i \left[\lambda_i \tilde{v}_i^i + \sum_{j \neq i} \lambda_j \alpha(i, j) \rho(y_b^i | \tilde{\mathbf{a}}^i) \right] - 1} \right)^{-1} \quad (9)$$

One way to contrast our approach with that of FLM (and FLT) 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_μ . By definition we must find a current action profile \mathbf{a} and continuation payoffs γ . 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 1.

Note that δ_μ is *decreasing* in μ . Since Condition 3 puts an absolute lower bound on μ and Condition 4 puts an absolute lower bound on δ_μ this means that (subject to the regularity constraint) there is a μ^* such that V_{μ^*} is the largest self-generating set (of this form) and δ_{μ^*} 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 4, for the two-player case, we do obtain a complete characterization of (efficient) PPE payoffs and we demonstrate the same phenomenon.

6. Perfect Public Equilibrium

Because every payoff in a self-generating set can be achieved in a PPE, Theorem 1 immediately provides sufficient conditions achieving (some) given target payoffs in perfect public equilibrium. In fact, we can provide an explicit algorithm for *computing* PPE strategies. A consequence of this algorithm is that (at least when action spaces are finite), the constructed PPE enjoys an interesting and potentially useful robustness property.

6.1. A Constructing Efficient Perfect Public Equilibria

Given the various parameters of the environment (game payoffs, information structure, discount factor) and of the problem (lower bound, target vector), the algorithm takes as input in period t the current continuation vector $v(t)$ and computes, for each player j , an indicator $d_j(v(t))$ defined as follows:

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

(Note that each player can compute every d_j from the current continuation vector $v(t)$ and the various parameters.) Having computed $d_j(v(t))$ for each j , the algorithm finds the player i^* whose indicator 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 uses the labeling $Y = \{y_g^{i^*}, y_b^{i^*}\}$ to compute continuation values for each signal in Y .

Theorem 2. *If the conditions in Theorem 1 are satisfied, then every payoff $v \in V_\mu$ can be achieved in a PPE. For $v \in V_\mu$, a PPE strategy profile that achieves v can be computed by the algorithm in Table 2*

Table 2: The algorithm used by each player.

Input: 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:

if $y^t = y_g^i$ **then**

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

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

for all $j \neq i$

if $y^t = y_b^i$ **then**

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

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

for all $j \neq i$

6.2. Robustness

A consequence of our constructive algorithm is that, for generic values of the parameters of the environment and of the problem and for as many periods as we specify, the strategies we identify are locally constant in these parameters. To make this precise, we assume for this subsection that action spaces A_i are finite. The parameters of the model are the utility mapping $U : \mathbf{A} \rightarrow \mathbb{R}^n$ and the probabilities $\rho(\cdot|\cdot) : Y \times \mathbf{A} \rightarrow [0, 1]$. Because the probabilities must sum to 1 and we require full support, the parameter space of the model is

$$\Omega = (R^n \times [0, 1])^{\mathbf{A}}$$

The parameters of the problem are the discount factor δ , the constraint vector μ and the target profile v^* ; because the target profile lies in a hyperplane, the parameter space for the particular problem is

$$\Theta = (0, 1) \times \mathbb{R}^n \times \mathbb{R}^{n-1}$$

Let $\Xi \subset \Omega \times \Theta$ be the subset of parameters that satisfy the Conditions of Theorem 1. For $\xi \in \Xi$, the algorithm generates an strategy profile

$$\sigma_\xi : \mathcal{H} \rightarrow \mathbf{A}$$

For $T \geq 0$ we write σ_ξ^T for the restriction of σ_ξ to the set \mathcal{H}^T of histories of length at most T .

Theorem 3. *For each $T \geq 0$ there is a subset $\Xi_T \subset \Xi$ that is closed and has measure 0 with the property that the mapping $\xi \rightarrow \sigma_\xi^T : \Xi \rightarrow \mathcal{H}^T$ is locally*

constant on the complement of Ξ_T .

In words: if ξ, ξ' are close together and neither lies in the proscribed small set of parameters Ξ_T , then the strategies $\sigma_\xi, \sigma_{\xi'}$ coincide for at least the first T periods.

6.3. Comparison with FLT

As we have commented in the Introduction, our approach provides a great deal of information about the efficient payoffs that can be achieved in PPE but because the sets V_μ are required to have a special form, it does not find all of them. Here we provide a simple example. We consider a 3×3 game. Each player chooses from the actions $\{l, m, h\}$: Player 1 chooses rows, Player 2 chooses columns, Player 3 chooses matrices; see Table 3. (Payoffs indicated by * are irrelevant so long as Assumptions 1,2 are satisfied; we could take $* = 0$ everywhere.) There are two signals y_g, y_b and the signal structure is

$$\rho(y_g|\mathbf{a}) = \begin{cases} 2/3 & \text{if } \mathbf{a} = (h, \ell, \ell) \text{ or any permutation} \\ 1/2 & \text{if } \mathbf{a} = (h, m, \ell) \text{ or any permutation} \\ 1/3 & \text{otherwise} \end{cases}$$

Note that $\tilde{\mathbf{a}}^1 = (h, \ell, \ell)$, $\tilde{\mathbf{a}}^2 = (\ell, h, \ell)$, $\tilde{\mathbf{a}}^3 = (\ell, \ell, h)$ and that $\tilde{v}^1 = (1, .5, 0)$, $\tilde{v}^2 = (0, 1, .5)$, $\tilde{v}^3 = (.5, 0, 1)$. Condition 3 implies that no regular V_μ can be a self-generating set (because we would have to have $\mu_i > .5$ for each i), so our approach does not find any PPE. However, applying the machinery of FLT shows that there is a discount factor $\hat{\delta} < 1$ for which the payoff vector $(.5, .5, .5)$ – indeed, any efficient payoff vector close to $(.5, .5, .5)$ – can be

Table 3: Payoff Matrices for the 3×3 Game; Player 3 Chooses ℓ , m , h (respectively)

	ℓ	m	h
ℓ	$(*, *, *)$	$(*, *, *)$	$(0, 1, 0.5)$
m	$(*, *, *)$	$(*, *, *)$	$(0.1, *, *)$
h	$(1, 0.5, 0)$	$(*, 0.55, *)$	$(0.2, 0.6, *)$

	ℓ	m	h
ℓ	$(*, *, *)$	$(*, *, *)$	$(*, *, 0.55)$
m	$(*, *, *)$	$(*, *, *)$	$(*, *, *)$
h	$(*, *, 0.1)$	$(*, *, *)$	$(*, *, *)$

	ℓ	m	h
ℓ	$(0.5, 0, 1)$	$(*, 0.1, *)$	$(*, 0.2, 0.6)$
m	$(0.55, *, *)$	$(*, *, *)$	$(*, *, *)$
h	$(0.6, *, 0.2)$	$(*, *, *)$	$(*, *, *)$

achieved in PPE.⁸ As noted in the Introduction, however, FLT provides no information as to what $\hat{\delta}$ must be nor does it construct PPE strategies.

7. Two Players

Theorem 1 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. 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

⁸Calculations available from the authors by request.

property that any target payoff in V that can be achieved as a PPE for *some* δ can already be achieved for *every* $\delta \geq \delta^*$.

Theorem 4. *Assume $N = 2$ (two players). Either*

- (i) *no target profile in V can be supported in a PPE for any $\delta < 1$ or*
- (ii) *there exist $\bar{\mu}_1, \bar{\mu}_2$ and a discount factor $\delta^* < 1$ such that if δ is any discount factor with $\delta^* \leq \delta < 1$ then the set of payoff vectors that can be supported in a PPE when the discount factor is δ is precisely*

$$E = \{v \in V : v_i \geq \bar{\mu}_i \text{ for } i = 1, 2\}$$

The proof yields explicit (messy) expressions for $\bar{\mu}_1, \bar{\mu}_2$ and δ^* .

8. Examples, Redux

In Section 3 we presented three examples to illustrate the model. We now return to these models to illustrate our analysis and conclusions.

Example 1 Because there are only two players, Theorem 4 applies. In this case it is easy to give explicit expressions for $\bar{\mu}_1, \bar{\mu}_2$ and for the threshold discount factor δ^* .⁹ In fact $\bar{\mu}_1 = \bar{\mu}_2 = q/(q-r)b$ and

$$\delta^* = \frac{1}{1 + \left(\frac{B - 2\frac{q}{q-r}b}{B + 2\frac{1-q}{q-r}b} \right)}$$

so that for every $\delta \geq \delta^*$ the set of efficient payoffs that can be achieved in PPE is exactly

$$E(\delta) = \{(v_1, v_2) : v_1 + v_2 = B; v_i \geq q/(q-r)b\}$$

We stress that the set of efficient equilibrium outcomes *does not* increase as $\delta \rightarrow 1$; as we noted in the Introduction, patience is rewarded but only up to a point. See Figure 5.

Example 2 As we have noted, the choice of a reporting rule has implications for the equilibria of the repeated game. In this case it is natural to consider two reporting rules $(Y^1, \psi^1), (Y^2, \psi^2)$ where

- $Y^1 = \{y_b, y_g\}; \psi^1(z_k) = y_b$ if $k = 0, \psi^1(z_k) = y_g$ if $k \neq 0$
- $Y^2 = Z; \psi^2(z_k) = z_k$ for all $k = 0, \dots, n$

⁹Calculations are available from the authors on request.

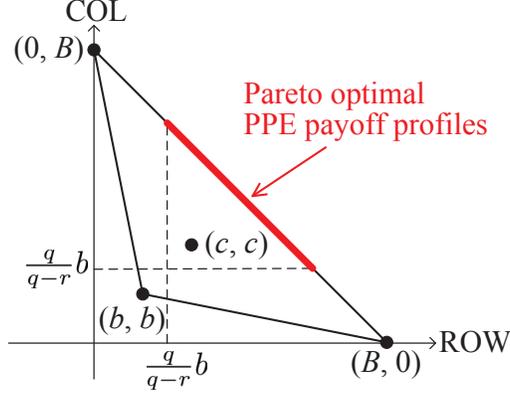


Figure 5: Efficient PPE Payoffs for the Modified Prisoners' Dilemma.

In the first rule (which is the one discussed in Section 3), the designer announces whether or not there has been a winner; in the second case the designer also announces the identity of the winner.

As we have noted earlier, in the reduced forms of the stage game, the *ex ante* expected utilities are given by

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

and, with the first reporting rule, the signal distribution is

$$\bar{p}(y_*|a) \begin{cases} 1 - \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } * = b \\ \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } * = g \end{cases}$$

With the second reporting rule, the signal distribution is

$$\widehat{\rho}(y_k|a) = \begin{cases} 1 - \sum_i a_i \left(\eta - \kappa \sum_{j \neq i} a_j \right)^+ & \text{if } k = 0 \\ a_k \left(\eta - \kappa \sum_{j \neq k} a_j \right)^+ & \text{if } k \neq 0 \end{cases}$$

To be specific, suppose there are 2 players. With the first reporting rule, there are two signals; with the second reporting rule there are three signals. The second reporting rule provides additional information to players and this additional information can be used to support more PPE. Suppose for instance that a strategy profile σ calls for $\tilde{\mathbf{a}}^1$ to be played after a particular history. If all the players follow σ then only player 1 exerts non-zero effort so only two outcomes can occur: either player 1 wins or no one wins. If player 2 deviates by exerting non-zero effort, a third outcome can occur: 2 wins. With either monitoring structure, it is possible for player 1 to detect (statistically) when player 2 has deviated, but with the second monitoring structure it will sometimes be the case that player 1 can be *certain* that player 2 has deviated. This additional information makes it possible to provide additional punishments for deviation and hence to support a larger set of efficient PPE. As Figure 6 shows, the difference matters: the second reporting rule always supports a larger set of efficient PPE; indeed, for some values of κ (which measures the strength of the interference) only the second reporting rule supports any efficient PPE at all.

More generally, consider a context with $n \geq 3$ players. If the identity of the winner is not announced, all deviations must be “punished” in the same way, but if the identity of the winner is announced, “punishments” can be tailored to the deviator, and hence can be more severe. If punishments can

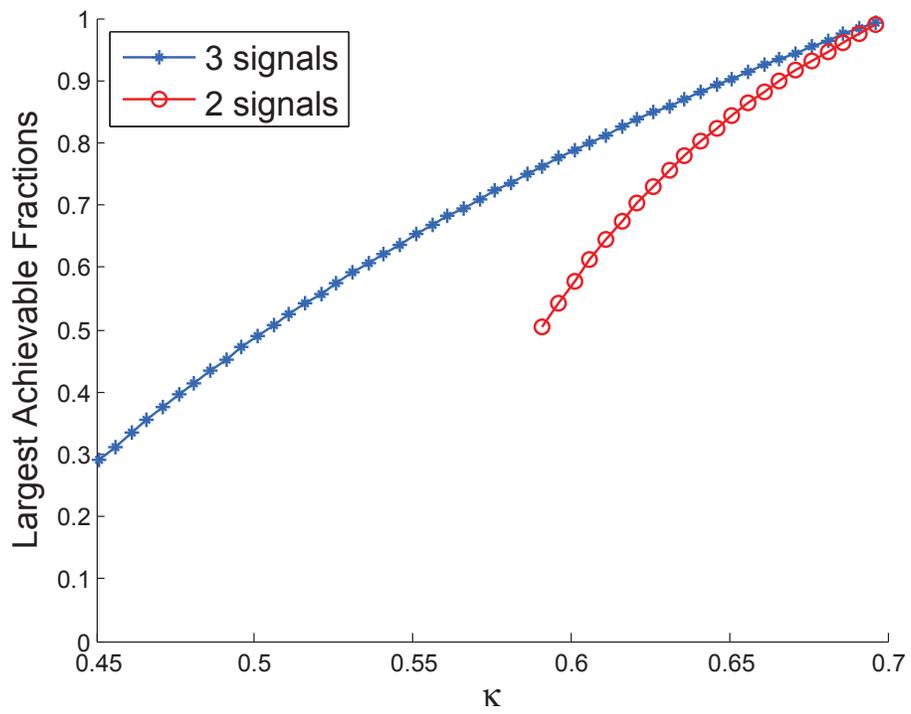


Figure 6: Largest Achievable Fraction $1 - \mu_i/\tilde{v}_i^i$ as a Function of κ .

be more severe it may be possible to sustain a wider range of PPE. Which reporting rule – hence which monitoring structure – should be chosen by the designer will depend on the tradeoff the designer makes between preserving privacy and sustaining a wider range of PPE.

Example 3 As we have suggested, the designer must choose from some (possible) measurement technologies. This choice involves a tension: more accurate measurement technologies will typically be more costly to employ. Hence the designer must trade-off the accuracy of the measurement technology against the cost of employing it. The designer must also choose a reporting rule, in this case a threshold d_0 . This choice also involves a tension, but of a different kind. Given the distribution of shocks and a choice of measurement technology, the choice of threshold affects the distribution of signals. How the designer chooses the distribution of signals depends on what the designer wishes to accomplish. For instance, given a fixed discount factor, the designer may wish to choose the threshold to maximize the range of long-run resource allocations that can be supported as PPE for the given discount factor. Alternatively, the designer may wish to minimize the discount factor for which *some* long-run resource allocation can be supported as a PPE.

To give some idea of the effect of these tradeoffs, we present numerical results for a special case of the Resource Sharing Game with 3 players, capacity $\chi = 1$ and $\bar{\varepsilon} = 0.3$. Because the game is symmetric it seems natural to consider symmetric sets of payoffs; so we consider sets of the form

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

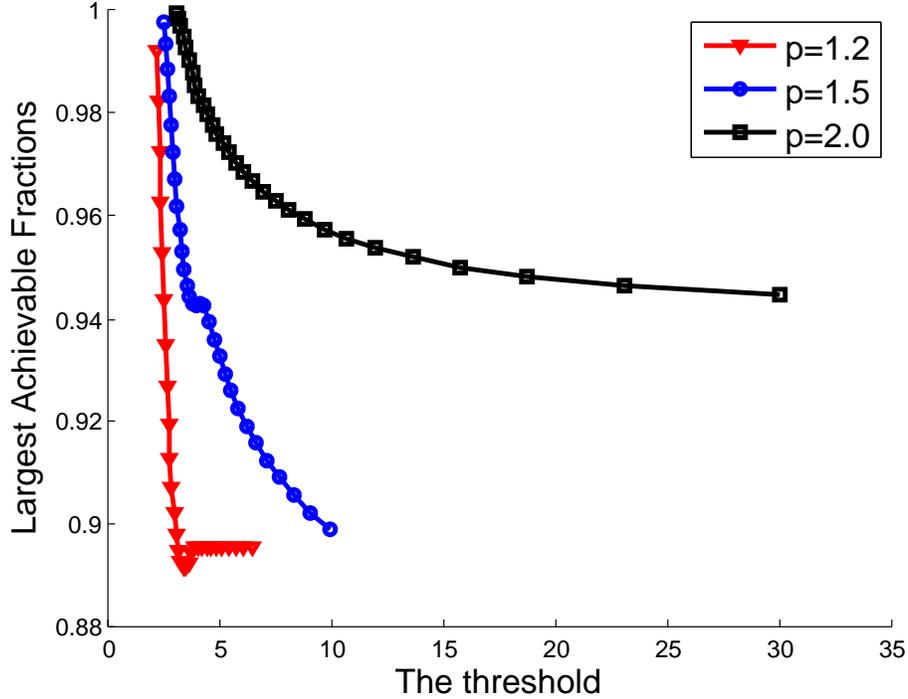


Figure 7: Largest Achievable Fraction $1 - \eta$ as a Function of Threshold d_0 .

where \tilde{v} is the utility of each player's most preferred action and $\eta \in [0, 1]$. Note that $1 - \eta$ represents the fraction of the entire efficient set V that is occupied by $V(\eta)$. A natural *desideratum* for the designer is to choose the threshold d_0 so that the fraction $1 - \eta$ is as large as possible; this maximizes opportunities for sharing. (As we have shown in Theorem 1, making η smaller also makes the required discount factor smaller, so the designer can simultaneously create more sharing opportunities for less patient players.) Figures 7 and 8 display (from simulations) the relationship between the threshold d_0 and the smallest η and smallest δ for different values of the exponent p .

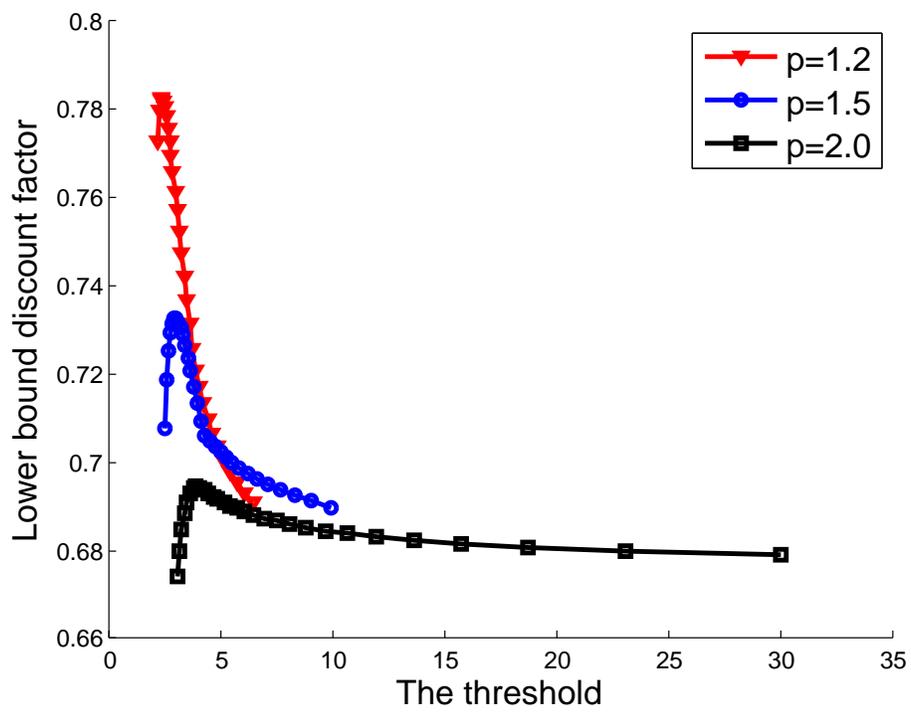


Figure 8: Smallest Achievable Discount Factor δ as a Function of Threshold d_0

9. Conclusion

This paper diverges from much of the familiar literature on repeated games with imperfect public monitoring in two directions. In analyzing the reduced form, we make different assumptions on the signal structure and obtain stronger conclusions about efficient PPE (bounds on the discount factor, explicitly constructive strategies). However, we also construct an elaborated form in which the information structure can be viewed as arising from the behavior of a strategic designer. Clearly there is much more to be done. Perhaps most obviously, it is clearly important to understand the extent to which the assumptions on the signal structure of the reduced form and on the geometry of the candidate self-generating sets V_μ can be relaxed. However, we think the elaborated form is of even more potential interest, especially for applications. As we have discussed in the Examples, the designer must decide what to observe and what to communicate to the players, and these choices will typically involve a trade-off between the cost of more accurate observation and communication on the one hand and the benefits of better information on the other hand. It seems natural to suppose that the costs and benefits – and hence the trade-offs – may be very different across environments. This seems a subject worthy of much study.

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, J. M., 1981. A new approach to performance-oriented flow control. *IEEE Transactions on Communications* 29 (4), 427–435.
- Blume, L. E., Zame, W. R., 1994. The algebraic geometry of perfect and sequential equilibrium. *Econometrica* 62 (4), 783–794.
- Bochnak, J., Coste, M., Roy, M.-F., 1998. *Real algebraic geometry*. Springer.
- Fudenberg, D., Levine, D. K., Maskin, E., 1994. The folk theorem with imperfect public information. *Econometrica* 62 (5), 997–1039.
- Fudenberg, D., Levine, D. K., Takahashi, S., 2007. Perfect public equilibrium when players are patient. *Games and Economic Behavior* 61 (1), 27 – 49.
- Green, E. J., Porter, R. H., 1984. Noncooperative collusion under imperfect price information. *Econometrica* 52 (1), 87–100.
- 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.
- Mailath, G., Samuelson, L., 2006. *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, Oxford, U.K.

Appendix

The proof of Proposition 1 is immediate and omitted.

Proof of Proposition 2 Fix an active player i and an inactive player j .

Set

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

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 σ 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 σ 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 σ ; because σ is a PPE, the incentive compatibility condition for all players j must obtain. Hence for all $a_j \in A_j$ we have

$$\begin{aligned} v_j &= (1 - \delta)u_j(\tilde{\mathbf{a}}^i) + \delta \sum_{y \in Y} \rho(y | \tilde{\mathbf{a}}^i) \gamma_j(y) \\ &\geq (1 - \delta)u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) + \delta \sum_{y \in Y} \rho(y | a_j, \tilde{\mathbf{a}}_{-j}^i) \gamma_j(y), \end{aligned} \quad (10)$$

Substituting probabilities for the good and bad signals yields

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

Rearranging yields

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

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 4) so

$$\left[\gamma_j(y_g^i) - \gamma_j(y_b^i) \right] \left[\frac{\delta}{1 - \delta} \right] \geq \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(y_b^i | a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)} \quad (11)$$

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

$$\left[\gamma_j(y_g^i) - \gamma_j(y_b^i) \right] \left[\frac{\delta}{1 - \delta} \right] \leq \frac{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - u_j(\tilde{\mathbf{a}}^i)}{\rho(y_b^i | a_j, \tilde{\mathbf{a}}_{-j}^i) - \rho(y_b^i | \tilde{\mathbf{a}}^i)} \quad (12)$$

Taking the sup over $a_j \in A(i, j)$ in (11) and the inf over $a_j \in B(i, j)$ in (12) yields $\alpha(i, j) \leq \beta(i, j)$ as desired. \square

Proof of Proposition 3 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, $u_i(\tilde{\mathbf{a}}^i) > u_i(a_i, \tilde{\mathbf{a}}_{-i}^i)$. With respect to probabilities, there are two possibilities.

If $\rho(y_b^i|a_i, \tilde{\mathbf{a}}_{-i}^i) \leq \rho(y_b^i|\tilde{\mathbf{a}}^i)$ then we immediately have

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

because the left-hand side is positive and the right-hand side is non-negative.

If $\rho(y_b^i|a_i, \tilde{\mathbf{a}}_{-i}^i) > \rho(y_b^i|\tilde{\mathbf{a}}^i)$ we proceed as follows.

We begin with (10) but now we apply it to the active user i , so that for all $a_i \in A_i$ we have

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

Rearranging yields

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

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

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

Applying the incentive compatibility constraints for the inactive users implies

that for each $a_j \in A(i, j)$ we have

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

In particular

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

and hence

$$\gamma_i(y_g^i) - \gamma_i(y_b^i) \leq -\frac{1}{\lambda_i} \left[\frac{1 - \delta}{\delta} \right] \left[\sum_{j \neq i} \lambda_j \alpha(i, j) \right] \leq 0$$

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

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

which is the desired result. \square

Proof of Theorem 1 Assume that V_μ is regular and not an extreme point, and is a self-generating set; we verify Conditions 1-4 in turn. Because V_μ is self-generating and not an extreme point, it cannot be a singleton and hence must contain an interior point of V . In order for such a point to be achieved in a PPE, every player must be active following some history, so Propositions 2 and 3 yield Conditions 1 and 2.

By assumption, for each $i \in N$ there is a payoff profile $\hat{v}^i \in V_\mu$ with the property that $\hat{v}_j^i = \mu_j$ for each $j \neq i$. Necessarily, \hat{v}^i is the unique such point

and $\hat{v}^i = \arg \max\{v_i : v \in V_\mu\}$. Because V lies in the hyperplane H we have

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

Because V_μ is self-generating, we can decompose \hat{v}^i :

$$\hat{v}^i = (1 - \delta)u(\tilde{\mathbf{a}}^k) + \delta \sum_y \rho(y|\tilde{\mathbf{a}}^k)\gamma(y) \quad (13)$$

for some $\tilde{\mathbf{a}}^k$. If $k \neq i$ then (because $V_\mu \neq \{\tilde{v}_k^k\}$) we must have $\mu_k < \tilde{v}_k^k$ which implies that $\gamma_k(y) < \mu_k$ for some y ; since continuation payoffs must lie in V_μ this is a contradiction. Hence in the decomposition (13) we must have $\tilde{\mathbf{a}}^k = \tilde{\mathbf{a}}^i$.

It is convenient to first establish the following inequality on μ_j on the way to establishing the bounds in Condition 3.

$$\mu_j > \max_{i \neq j} \tilde{v}_j^i \text{ for all } j \in N$$

To see this, suppose to the contrary that there exists a i, j such that $\mu_j \leq \tilde{v}_j^i$. Consider i 's preferred payoff profile \hat{v}^i in V_μ . Because decomposing \hat{v}^i requires that we use $\tilde{\mathbf{a}}^i$, it follows that

$$\mu_j = (1 - \delta) \cdot \tilde{v}_j^i + \delta \cdot \sum_y \rho(y|\tilde{\mathbf{a}}^i)\gamma_j(y)$$

If $\mu_j < \tilde{v}_j^i$ then $\sum_{y \in Y} \rho(y|\tilde{\mathbf{a}}^i)\gamma_i(y) < \mu_j$ and so $\gamma_j(y) < \mu_j$ for some y . This contradicts that fact that $\gamma(y) \in V_\mu$. If $\mu_j = \tilde{v}_j^i$, we must have

$\sum_y \rho(y|\tilde{\mathbf{a}}^i) \gamma_j(y) = \mu_j$. Since $\gamma_j(y) \geq \mu_j$ for all y , we must have $\gamma_j(y_g^i) = \gamma_j(y_b^i) = \mu_j$. By assumption, player j has a currently profitable deviation a_j so that $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$, which implies that the continuation payoff $\gamma_j(y_g^i) = \gamma_j(y_b^i) = \mu_j$ cannot satisfy the incentive compatibility constraints. Hence, we must have $\mu_j > \tilde{v}_j^i$ as asserted.

With all this in hand we derive Condition 3. 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

$$\begin{aligned} v_j &= (1 - \delta) \cdot \tilde{v}_j^i + \delta \cdot \sum_{y \in Y} \rho(y|\tilde{\mathbf{a}}^i) \cdot \gamma_j(y) \\ &\geq (1 - \delta) \cdot u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) + \delta \cdot \sum_{y \in Y} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(y). \end{aligned} \tag{14}$$

From the equality constraint in (14), we can solve for the discount factor δ 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 δ in the inequality of (14). 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

$$v_j \leq \sum_{y \in Y} \gamma_j(y) \left[\left(1 - \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right) \rho(y|\tilde{\mathbf{a}}^i) + \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \right] \quad (15)$$

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

$$v_j \geq \sum_{y \in Y} \gamma_j(y) \left[\left(1 - \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \right) \rho(y|\tilde{\mathbf{a}}^i) + \frac{v_j - \tilde{v}_j^i}{u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) - \tilde{v}_j^i} \rho(y|a_j, \tilde{\mathbf{a}}_{-j}^i) \right] \quad (16)$$

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

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

According to (15), if $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > \tilde{v}_j^i$ then

$$c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(y_g^i) + \left[1 - c_{ij}(a_j, \tilde{\mathbf{a}}_{-j}^i) \right] \gamma_j(y_b^i) \leq v_j \quad (17)$$

Since $\gamma_j(y_g^i) > \gamma_j(y_b^i)$, this is true if and only if

$$\kappa_{ij}^+ \cdot \gamma_j(y_g^i) + (1 - \kappa_{ij}^+) \cdot \gamma_j(y_b^i) \leq v_j, \quad (18)$$

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 (17) 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 (18). If (18) is satisfied, then the inequalities (17) are satisfied for all a_j such that $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$ because $\gamma_j(y_g^i) > \gamma_j(y_b^i)$ 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 (17) are satisfied for all a_j such that $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) > u_j(\tilde{\mathbf{a}}^i)$ and (18) were violated, so that $\kappa_{ij}^+ \cdot \gamma_j(y_g^i) + (1 - \kappa_{ij}^+) \cdot \gamma_j(y_b^i) > v_j$, then we can find a $\kappa'_{ij} < \kappa_{ij}^+$ such that $\kappa'_{ij} \cdot \gamma_j(y_g^i) + (1 - \kappa'_{ij}) \cdot \gamma_j(y_b^i) > v_j$. 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 $c_{ij}(a'_j, \tilde{\mathbf{a}}_{-j}^i) \cdot \gamma_j(y_g^i) + (1 - c_{ij}(a'_j, \tilde{\mathbf{a}}_{-j}^i)) \cdot \gamma_j(y_b^i) > v_j$. This contradicts the fact that the inequalities (18) 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 (16), for all a_j such that $u_j(a_j, \tilde{\mathbf{a}}_{-j}^i) < \tilde{v}_j^i$, we must have

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

Since $\gamma_j(y_g^i) > \gamma_j(y_b^i)$, the above requirement is fulfilled if and only if

$$\kappa_{ij}^- \cdot \gamma_j(y_g^i) + (1 - \kappa_{ij}^-) \cdot \gamma_j(y_b^i) \geq v_j,$$

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 decom-

position (14) for user $j \neq i$ can be simplified as:

$$\begin{aligned}
\rho(y_g^i | \tilde{\mathbf{a}}^i) \cdot \gamma_j(y_g^i) + [1 - \rho(y_g^i | \tilde{\mathbf{a}}^i)] \gamma_j(y_b^i) &= \tilde{v}_j^i + \frac{v_j - \tilde{v}_j^i}{\delta} \\
\kappa_{ij}^+ \gamma_j(y_g^i) + (1 - \kappa_{ij}^+) \cdot \gamma_j(y_b^i) &\leq v_j \\
\kappa_{ij}^- \gamma_j(y_g^i) + (1 - \kappa_{ij}^-) \cdot \gamma_j(y_b^i) &\geq v_j
\end{aligned} \tag{19}$$

Keep in mind that the various continuation values γ and the expressions $\kappa_{ij}^+, \kappa_{ij}^-$ depend on v_j ; where necessary we write the dependence explicitly. Note that there could be many $\gamma_j(y_g^i)$ and $\gamma_j(y_b^i)$ that satisfy (19). For a given discount factor δ , we call all the continuation payoffs that satisfy (19) *feasible* – but whether particular continuation values lie in V_μ 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 again at player i 's preferred payoff profile \hat{v}^i in V_μ , which is necessarily decomposed by $\tilde{\mathbf{a}}^i$. We look at the following constraint for player $j \neq i$ in (19):

$$\kappa_{ij}^+ \gamma_j(y_g^i) + (1 - \kappa_{ij}^+) \gamma_j(y_b^i) \leq \mu_j.$$

Suppose that $\kappa_{ij}^+(\mu_j) > 0$. Since player j has a currently profitable deviation from $\tilde{\mathbf{a}}^i$, we must set $\gamma_j(y_g^i) > \gamma_j(y_b^i)$. Then to satisfy the above inequality, we must have $\gamma_j(y_b^i) < \mu_j$. In other words, when $\kappa_{ij}^+(\mu_j) > 0$, all the feasible continuation payoffs of player j must be outside V_μ . This contradicts the fact that V_μ 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

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

This provides a lower bound on μ_j :

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

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

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

which is Condition 3.

Now we derive Condition 4 (the necessary condition on the discount factor). The minimum discount factor $\underline{\delta}_\mu$ required for V_μ to be a self-generating set solves the optimization problem

$$\underline{\delta}_\mu = \max_{v \in V_\mu} \delta \quad \text{subject to } v \in \mathcal{B}(V_\mu; \delta)$$

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

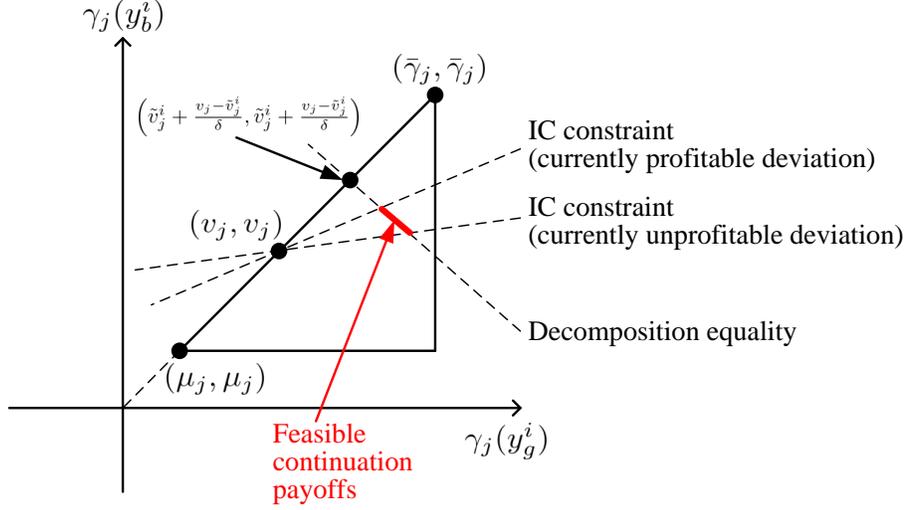


Figure 9: Illustrations of the feasible continuation payoffs when $\kappa_{ij}^+ \leq 0$. $\bar{\gamma}_j = \frac{1}{\lambda_j} \left(1 - \sum_{k \neq j} \lambda_k \mu_k\right)$.

mization problem can be reformulated as

$$\underline{\delta}_\mu = \max_{v \in V_\mu} \min_{i \in N} \delta \quad \text{subject to } v \in \mathcal{B}(V_\mu; \delta, \tilde{\mathbf{a}}^i). \quad (20)$$

To solve the optimization problem (20), 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 $\gamma_j(y_g^i)$ and $\gamma_j(y_b^i)$ must satisfy (19). There are many $\gamma_j(y_g^i)$ and $\gamma_j(y_b^i)$ that satisfy (19). In Fig. 9, we show the feasible continuation payoffs that satisfy (19) 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 $\rho(y_g^i | \tilde{\mathbf{a}}^i) \cdot \gamma_j(y_g^i) + (1 - \rho(y_g^i | \tilde{\mathbf{a}}^i)) \cdot \gamma_j(y_b^i) = \tilde{v}_j + \frac{v_j - \tilde{v}_j}{\delta}$, and is bounded by the IC constraint on currently profitable deviations $\kappa_{ij}^+ \cdot \gamma_j(y_g^i) + (1 - \kappa_{ij}^+) \cdot$

$\gamma_j(y_b^i) \leq v_j$ and the IC constraint on currently unprofitable deviations $\kappa_{ij}^- \cdot \gamma_j(y_g^i) + (1 - \kappa_{ij}^-) \cdot \gamma_j(y_b^i) \geq v_j$. Among all the feasible continuation payoffs, denoted $\gamma'(y)$, we choose the one, denoted $\gamma^*(y)$, such that for all $j \neq i$, $\gamma_j^*(y_g^i)$ and $\gamma_j^*(y_b^i)$ make the IC constraint on currently profitable deviations in (19) binding. This is because under the same discount factor δ , 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. 9, the continuation payoff $\gamma_j^*(y)$ that makes the IC constraint binding has the smallest $\gamma_j^*(y_g^i) = \min \gamma_j'(y_g^i)$ and the largest $\gamma_j^*(y_b^i) = \max \gamma_j'(y_b^i)$. Formally we establish the following Lemma.

Lemma 1. *Fix a payoff profile v and a discount factor δ . Suppose that v is decomposed by $\tilde{\alpha}^i$. If there are any feasible continuation payoffs $\gamma'(y_g^i) \in V_\mu$ and $\gamma'(y_b^i) \in V_\mu$ that satisfy (19) for all $j \neq i$, there there exist feasible continuation payoffs $\gamma^*(y_g^i) \in V_\mu$ and $\gamma^*(y_b^i) \in V_\mu$ such that the IC constraint on currently profitable deviations in (19) is binding for all $j \neq i$.*

Proof. Given feasible continuation payoffs $\gamma'(y_g^i) \in V_\mu$ and $\gamma'(y_b^i) \in V_\mu$, we construct $\gamma^*(y_g^i) \in V_\mu$ and $\gamma^*(y_b^i) \in V_\mu$ that are feasible and make the IC constraint on currently profitable deviations in (19) binding for all $j \neq i$.

Specifically, we set $\gamma_j^*(y_g^i)$ and $\gamma_j^*(y_b^i)$ such that the IC constraint on currently profitable deviations in (19) is binding. Such $\gamma_j^*(y_g^i)$ and $\gamma_j^*(y_b^i)$ have the following property: $\gamma_j^*(y_g^i) \leq \gamma_j'(y_g^i)$ and $\gamma_j^*(y_b^i) \geq \gamma_j'(y_b^i)$ for all $\gamma_j'(y_g^i)$ and $\gamma_j'(y_b^i)$ that satisfy (19). We prove this property by contradiction. Suppose that there exist $\gamma_j'(y_g^i)$ and $\gamma_j'(y_b^i)$ that satisfy (19) and $\gamma_j'(y_g^i) = \gamma_j^*(y_g^i) - \Delta$

with $\Delta > 0$. Based on the decomposition equality, we have

$$\gamma'_j(y_b^i) = \gamma_j^*(y_b^i) + \left(\frac{\rho(y_g^i | \tilde{\mathbf{a}}^i)}{1 - \rho(y_g^i | \tilde{\mathbf{a}}^i)} \right) \Delta$$

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

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

where the last inequality results from $\kappa_{ij}^+ \leq 0$. Hence, we have $\gamma_j^*(y_g^i) \leq \gamma'_j(y_g^i)$ and $\gamma_j^*(y_b^i) \geq \gamma'_j(y_b^i)$ for all $\gamma'_j(y_g^i)$ and $\gamma'_j(y_b^i)$ that satisfy (19).

Next, we prove that if $\gamma'(y) \in V_\mu$, then $\gamma^*(y) \in V_\mu$. To prove $\gamma^*(y) \in V_\mu$, we need to show that $\gamma_j^*(y_g^i) \geq \mu_j$ and $\gamma_j^*(y_b^i) \geq \mu_j$ for all $j \in N$. For $j \neq i$, we have $\gamma_j^*(y_g^i) \geq \gamma_j^*(y_b^i) \geq \gamma'_j(y_b^i) \geq \mu_j$. For i , we have

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

This proves the lemma. \square

Using this Lemma, we can calculate the continuation payoffs of the inac-

tive player $j \neq i$:

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

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

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

Hence, the constraint $v \in \mathcal{B}(V_\mu; \delta, \tilde{\mathbf{a}}^i)$ on discount factor δ 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 $\gamma_i(y_g^i) < \gamma_i(y_b^i)$, we need

$\gamma_i(y_g^i) \geq \mu_i$, which leads to

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

Hence, the optimization problem (20) is equivalent to

$$\underline{\delta}(\mu) = \max_{v \in V_\mu} \min_{i \in \mathcal{N}} x_i(v) \quad (21)$$

where

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

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

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

Then we have

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

from which it follows that

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

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

It remains to show that these necessary Conditions are also sufficient, which is accomplished in the proof of Theorem 2. This completes the proof of Theorem 1. \square

Proof of Theorem 2 In view of the results of APS, it suffices to show that the algorithm yields a decomposition of each target vector $v(t) \in V_\mu$. The algebra in the proofs of Propositions 2 and 3 shows that Conditions 1 and 2 guarantee that the incentive compatibility constraints are satisfied for the inactive and active players. The algebra in the proof of Theorem 1 shows that Conditions 3 and 4 taken together guarantee that the continuation payoff $\gamma(y)$ belongs to V_μ for each $y \in Y$. \square

Proof of Theorem 3 Given a parameter ξ , the algorithm uses the target vector as the continuation value to compute N indicators; let $\Xi(0)$ be the set of parameters for which no two of these indicators are equal. For each parameter in $\Xi(0)$, the algorithm computes continuation values following the good signal and the bad signal and then uses each of these continuation values to compute N indicators; let $\Xi(1) \subset \Xi(0)$ be the set of parameters

for which no two of these indicators are equal. Proceeding by induction, we define a decreasing sequence of sets $\Xi(0) \supset \Xi(1) \supset \dots \supset \Xi(T)$; let Ξ_T be the complement of $\Xi(T)$. Notice that the indicators are continuous functions of the parameters so the ordering of the indicators is locally constant provided no two indicators are equal. Hence for each $\xi \in \Xi(T) = \Xi \setminus \Xi(T)$ then there is a small open neighborhood Z of ξ so that if $\xi' \in Z$ then the strategies $\sigma_{\xi'}, \sigma_{\xi}$ generate the same ordering of indicators in each of the first T periods. In particular, $\sigma_{\xi'}(h) = \sigma_{\xi}(h)$ for each history $h \in \mathcal{H}^T$; that is, $\xi \rightarrow \sigma_{\xi}^T$ is locally constant on the complement of Ξ_T . It remains only to show that Ξ_T is closed and has measure 0. In fact, Ξ_T is a finite union of lower-dimensional submanifolds; this is a consequence of general facts about semi-algebraic sets and the observation that all the indicators are continuous semi-algebraic functions of the parameters, no two of which coincide on any open set. See Bochnak, Coste, and Roy (1998), Blume and Zame (1994).

Proof of Theorem 4 Propositions 2, 3 show that Conditions 1, 2 are necessary conditions for the existence of an efficient PPE for *any* discount factor.. Suppose therefore that Conditions 1,2 are satisfied. It is easily checked that the definitions of $\bar{\mu}_1, \bar{\mu}_2$ guarantee that Condition 3 of Theorem 1 are satisfied. Finally, if $\delta \geq \delta^*$ then Condition 4 of Theorem 1 is also satisfied. It follows from Theorem 1 that for each $\delta \geq \delta^*$, $V_{\bar{\mu}}$ is a self-generating set, so every target vector in $V_{\bar{\mu}}$ can be achieved in a PPE. Hence $E(\delta) \supset V_{\bar{\mu}}$ for every $\delta \in [\delta^*, 1)$. To see that $V_{\bar{\mu}} = E(\delta)$ for every $\delta \in [\delta^*, 1)$, simply note that for each δ the set $E(\delta)$ is closed and convex, hence an interval, hence of the form V_{μ} for some μ . However, Condition 3 of Theorem 1 guarantees that $\mu \geq \bar{\mu}$ which completes the proof. \square