# Repeated Games, Part II: Imperfect Public Monitoring

Lecture Notes, Yale 2015, Johannes Hörner

June 27, 2016

## I An Example

This example follows Abreu, Milgrom and Pearce (1991). Consider the problem of a team of 5 workers, or players, who are working together on a project forever. We shall consider two variants of this infinitely repeated game, one in which their efforts reduce the probability of unfortunate events, and one in which it increases the probability of desirable ones. In both cases, denote each player's individual effort by $e_i \in [0,1]$, $i = 1, \ldots, 5$. Period length is $\Delta > 0$, and the cost of effort over some interval is $(e_i + e_i^2)\,\Delta/2$. At the beginning of each period, each player independently chooses his effort level, and his choice is not observed by the other players. Players all discount payoffs at rate $\delta := e^{-r\Delta}$, for some $r > 0$. For concreteness, let us pick $e^{-r} = 1/10$. Throughout, we assume $\Delta < 1/6$.

### A First Variant: The "Bad News" Case

In the first variant, effort reduces the probability of an accident in each period. Each accident yields a (dis)utility of $-1$ whenever it occurs (the corresponding utility if there is no accident is normalized to 0), and its probability in a period is given by

$$\left(6 - \sum_j e_j\right)\Delta.$$

Accidents are the only events that are observed. Player $i$'s reward in a period is then given by

$$-\left(6 - \sum_j e_j + \frac{e_i + e_i^2}{2}\right)\Delta.$$

In the one-shot game, each player maximizes this reward by choosing $e_i = 1/2$, which yields a reward of $-(31/8)\Delta$. Note that the social optimum involves all players putting in effort $e_i = 1$, for a reward of $-2\Delta$.

What is the best equilibrium in the repeated game? Note that, because of the quadratic cost, players will never find it optimal to randomize.[1] Further, because the only way to punish players is by increasing the probability of accidents, which affects all players equally, we can expect equilibria to be symmetric, *i.e.*, $e_i = e$.[2]

So let us characterize the best symmetric equilibrium, under the assumption that players have access to a public randomization device. Let us denote by $v^H$ the payoff of the best symmetric equilibrium, and let us also write $v^L$ for the worst (symmetric) equilibrium payoff. We refer to $H$ and $L$ as states. Achieving $v^H$ requires specifying an action $e^H$ to be played in the initial period, as well as a continuation payoff from the second period onward according to whether there is an accident or no accident: $w_A^H, w_N^H$. Similarly, achieving $v^L$ calls for an action $e^L$ to be played, followed by continuation payoffs $w_A^L, w_N^L$. Because players could implement the continuation equilibrium immediately, it must be that

$$w_A^H, w_N^H, w_A^L, w_N^L \in \left[v^L, v^H\right]. \tag{1}$$

Furthermore, $e^H, e^L$ must be Nash equilibria, with payoffs $v^H, v^L$, of the one-shot game with payoff functions

$$\begin{aligned}
&\Delta\left(6 - 4e^k - e_i^k\right)\left[(1-\delta)\left(-1 - \Delta\frac{e_i^k + (e_i^k)^2}{2}\right) + \delta w_A^k\right] \\
&+ \left(1 - \Delta\left(6 - 4e^k - e_i^k\right)\right)\left[-(1-\delta)\Delta\frac{e_i^k + (e_i^k)^2}{2} + \delta w_N^k\right],
\end{aligned} \tag{2}$$

with $k = H$ and $L$. Clearly, the payoff $v^H$ (resp. $v^L$) is increasing in $e^H$ (resp., $e^L$),[3] and so, to maximize the gap between $v^H$ and $v^L$ (so as to maximize incentives), we want to choose the largest (resp. smallest) effort consistent with (2), subject to (1). Note that, given our randomization device, we can achieve any payoff $w$ in the range $\left[v^L, v^H\right]$ simply by using the randomization device to coordinate on $v^L$ and $v^H$ with probability $\left(v^H - w\right)/\left(v^H - v^L\right)$ and $\left(w - v^L\right)/\left(v^H - v^L\right)$, respectively. Therefore, players can restrict attention to equilibria in which, after any history,

[1] That is, if a player were indifferent between two distinct effort levels, he would strictly prefer the expected effort level, as its cost would be lower than the expected cost.

[2] Proving this is the case requires a little work, though.

[3] Recall that first-best effort is 1.

and as a function of the public randomization device, either $e^H$ or $e^L$ is played. So define

$$q_A^H := \frac{v^H - w_A^H}{v^H - v^L}, \quad q_N^H := \frac{v^H - w_N^H}{v^H - v^L},$$

as well as

$$q_A^L := \frac{w_A^L - v^L}{v^H - v^L}, \quad q_N^L := \frac{w_N^L - v^L}{v^H - v^L},$$

as the probabilities of switching states, as a function of the initial state ($L$ or $H$) and the occurrence or not of an accident. It is not hard to see that, to maximize (resp., minimize) $v^H$ (resp., $v^L$), we must set

$$q_N^H = 0, \quad q_N^L = 0;$$

that is, to encourage effort in state $H$, it is best to avoid all punishment if an accident is avoided. Similarly, to discourage effort in state $L$, we should only consider switching back to state $H$ if an accident is observed. The switching probabilities can then be found by taking first-order conditions with respect to effort $e_i^j$ in (2) and simplifying:[4]

$$(1-\delta)\left(e^H - \frac{1}{2}\right) = \delta q_A^H \left(v^H - v^L\right), \quad (1-\delta)\left(\frac{1}{2} - e^L\right) = \delta q_A^L \left(v^H - v^L\right).$$

Solving for the switching probabilities, and plugging back into (2) gives

$$v^k = \left((9/2)\left(e^k\right)^2 - 4e^k - 3\right)\Delta.$$

So it is best to set $e^H = 1$, $e^L = 4/9$, and these satisfy (1) if $\Delta$ is small and $\delta$ is high enough (more precisely, if $\Delta \leq .158$, given $e^{-r} = 1/10$ ). This gives us

$$v^H = -5\Delta/2, \quad v^L = -35\Delta/9$$

Note that $v^H$ is independent of $r$. We picked $e^{-r} = 1/10$, but nothing changes if players are more patient. In particular, we do not get efficiency as $\delta \to 1$.

[4]These must hold even if effort is extreme, because otherwise one could decrease the probability of "punishment" (or increase the probability of "reward") and still get the same extreme effort, while increasing (or decreasing) $v^H$ (resp. $v^L$).

## B Second Variant: The "Good News" Case

In this variant, effort increases the probability of a desirable event —say, a sale, worth 1 to each player. Hence, let us assume that the probability of a sale is given by:[5]

$$\left(1+\sum_j e_j\right)\Delta.$$

The cost of effort is the same as before. We can analyze the game as before, solving for payoffs

$$w_S^H, w_N^H, w_S^L, w_N^L \in \left[v^L, v^H\right],$$

according to whether a sales occurs or not. Subject to this condition, the effort levels $e^H, e^L$ must be Nash equilibria, with payoffs $v^H, v^L$, of the one–shot game with payoff functions

$$\begin{aligned}&\left(1+4e^k+e_i^k\right)\Delta\left[(1-\delta)\left(1-\Delta\frac{e_i^k+\left(e_i^k\right)^2}{2}\right)+\delta w_S^k\right]\\&+\left(1-\left(1+4e^k+e_i^k\right)\Delta\right)\left[-(1-\delta)\Delta\frac{e_i^k+\left(e_i^k\right)^2}{2}+\delta w_N^k\right],\end{aligned}$$

with $k=H$ and $L$. This can be analyzed as before. The main difference is that the probabilities of switching after sales should, quite intuitively, be set equal to zero. But the consequences are huge: as you can easily check, the unique solution is to set $e^L=e^H=1/2$, which is the Nash equilibrium of the one-shot game: no collusion can be sustained, independently of patience.

## C Conclusions

We can draw several implications from this example:

1. Recursive methods appear to be applicable to games with imperfect monitoring. The next section will formalize the heuristic methods used here. This is the topic of Section II.

2. The folk theorem need not hold once monitoring is imperfect. In this example, in fact, the highest payoff is bounded away from the efficient payoff, no matter how patient players are. If there are sufficient conditions for the folk theorem, our example must fail those. It

[5] This keeps the range of probabilities the same as in the first variant.

is natural to wonder, then, what these sufficient conditions are, and this will be the topic of Section IV.

3. When the folk theorem fails, details matter a great deal. A characterization of the equilibrium payoff set for $\delta \to 1$ will be provided in Section III. In our example, the key distinction is whether the rare event is good news or bad news. What is the intuition behind the different results across variants? The key lies in the informativess of the bad news signal, which is the one that must trigger a punishment. In the first example, accidents are informative, as the likelihood ratio

$$\frac{\left(6-\sum_j e_j\right)\Delta}{\left(6-\sum_{j\neq i} e_j - e_i'\right)\Delta}$$

is independent of $\Delta$ and sensitive to $e_i'$: accidents carry information about actions taken by the players. In the second example, in contrast, the likelihood ratio

$$\frac{1-\left(1+\sum_j e_j\right)\Delta}{1-\left(1+\sum_j e_{j\neq i} + e_i'\right)\Delta}$$

converges to 1, independently of $e_i'$, as $\Delta \to 0$: bad news signal are no longer informative, and it is no longer possible to use them to collude effectively.

# II Recursive Methods

## A Notations

Attention is restricted throughout to infinitely repeated games. A repeated game with imperfect public monitoring specifies, in addition to the set of players $N = \{1, \ldots, n\}$, and action profiles $A$, a set of signals $Y$ (finite), and, for each action profile $a \in A$, a distribution $\pi(\cdot \mid a)$ on $Y$, the **monitoring structure**. The interpretation is straightforward: as a function of the action profile $a$ played in a given period, the signal $y \in Y$ is drawn according to the distribution $\pi(\cdot \mid a)$. Actions are not observed by other players; on the other hand, the signal $y$ is publicly observed. We write $\pi(y \mid \alpha) = \sum_a \alpha(a)\pi(y \mid a)$ for the distribution of signals induced by a mixed action $\alpha \in \triangle A$.

For consistency, rewards are usually first defined as maps

$$g_i : A_i \times Y \to \mathbb{R},$$

so that a player's realized reward $g_i\left(a_i, y\right)$ carries no more information than what he already knows, or observes, namely $a_i$ and $y$. Given some action profile $a$, player $i$'s expected reward is then

$$u_i\left(a\right) := \sum_{y \in Y} g_i\left(a_i, y\right) \pi\left(y \mid a\right),$$

whose average discounted sum he seeks to maximize. It has been customary to use the function $u$ as the primitive of the repeated game, rather than $g$. In this fashion, fixing $u$, we can examine how the quality of the monitoring affects the equilibrium payoff set without having to worry about how the change in monitoring affects the set of feasible payoffs, as it "mechanically" would if we were to take $g$ as a primitive.

Therefore, we shall take $u$ as a primitive, and ignore $g$ from now on, but it is important to keep in mind that players cannot infer anything from $u_i\left(a\right)$ beyond what they already know, $a_i$ and $y$.

A repeated game with imperfect public monitoring, then, is a collection $\left(N, A, Y, \pi\left(\cdot \mid a\right)_{a \in A}, u\right)$, along with a discount factor $\delta$. Note that, up to period $t$, player $i$ has observed an element of $H_i^t := \left(A_i \times Y\right)^t$, corresponding to the actions he has played, and the public signals he has observed.[6] This is the set of **private histories** $h_i^t$. Players share some information, namely the sequence of public signals, or **public history** $h^t \in H^t := Y^t$. Perfect monitoring is the special case in which $Y = A$, and $\pi\left(y \mid a\right) = 1$ iff $y = a$, so that action profiles are perfectly observed. Of course, our interest primarily lies in the case in which the monitoring is not perfect, though everything we shall prove applies to perfect monitoring as well.

# B Definitions

What we are trying to achieve here is a "simple," "recursive" representation of the set of equilibrium payoffs. We shall cheat, and define jointly the solution concept, a refinement of sequential equilibrium, and the recursive representation, which led to our choice.

**Definition 1** *The strategy $\sigma_i$ is* ***public*** *if it is measurable with respect to the public history: for*

[6] If a public randomization device is assumed, it is understood that each player also has observed its realizations in all previous periods.

*all $t$ and sequences $(y^s, a_i^s, \hat{a}_i^s)_{s=0}^t$,*

$$\sigma_i\left(a_i^1 y^1, a_i^2 y^2, \ldots, a_i^t y^t\right) = \sigma_i\left(\hat{a}_i^1 y^1, \hat{a}_i^2 y^2, \ldots, \hat{a}_i^t y^t\right).$$

If players $-i$ use public strategies, they disregard their private information going forward. Therefore, player $i$ has nothing to gain from conditioning on his private information either, and he has a best-reply that is public as well. This implies that public strategies are closed under best-replies.

**Definition 2** *The strategy profile $\sigma$ is a **public perfect equilibrium**, or **PPE**, if, for all $i$, $\sigma_i$ is public, and for all public histories $h^t$, $\sigma|_{h^t}$ is a Nash equilibrium of the repeated game.*

This solution concept is a natural extension of subgame-perfection to imperfect monitoring: if players' strategies only condition on public events, we require that they are Nash equilibria conditional on any such event. Clearly, PPE are sequential equilibria. However, there are sequential equilibria that are not PPE, and there are well-known examples of repeated games in which, as $\delta \to 1$, efficient payoffs can be approximated by sequential equilibria, but not by PPE: the power of statistical tests to detect deviations can be improved by using all information a player has available, which includes his own privately observed actions.[7]

We let $E_\delta$ denote the set of PPE payoffs, given the repeated game and the discount factor $\delta$. The main benefit of this solution concept is that the set of PPE payoffs is independent of the public history: of course, which PPE is selected as a continuation strategy profile depends on the public history, in general, but the set of PPE to select from does not. This is not true for sequential equilibria, as private histories provide private correlation devices whose structure depends, among others, on the period considered.

We now turn to the central tools for studying $E_\delta$, introduced by Shapley (1953), Mertens & Parthasarathy (1986) and Abreu, Pearce and Stacchetti (1990, hereafer referred to as APS). Recall that, if $W$ and $Y$ are sets, $W^Y$ is the set of functions from $Y$ to $W$.

**Definition 3** *Given $W \subset \mathbb{R}^n$, $\alpha \in \times_i \triangle A_i$ is **enforceable** on $W$ if there exists $w \in W^Y$ such that $\alpha$ is a Nash equilibrium of the game $\Gamma_\delta(w)$ with action sets $A_i$ and payoff function*

$$(1-\delta)\, u(\cdot) + \delta \sum_y \pi(y \mid \cdot)\, w(y). \tag{3}$$

[7] Kreps and Wilson (1982) define sequential equilibrium for finite games only. Throughout these notes, the definition is extended to infinitely repeated games by equipping both the set of strategies and the set of systems of beliefs with the uniform topology of uniform convergence over information sets.

The function $w$ **enforces** $\alpha$. The equilibrium payoff vector

$$v = (1-\delta)\, u(\alpha) + \delta \sum_y \pi(y \mid \alpha)\, w(y)$$

is **decomposed** by $(\alpha, w)$ on $W$, and it is **decomposable** on $W$ if there exists such a pair $(\alpha, w) \in \times_i \triangle A_i \times W^Y$.

Let $\mathcal{P}(\mathbb{R}^n)$ denote the set of all subsets of $\mathbb{R}^n$. We define the map

$$\begin{aligned} B &: \mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(\mathbb{R}^n) \\ & W \longmapsto B(W) = \{v \in \mathbb{R}^n : v \text{ is decomposable on } W\}. \end{aligned}$$

These are all the payoffs that can be obtained by decomposition, by using "continuation" payoffs from $W$ only. We shall also write $B_\delta$ rather than $B$ whenever convenient. Here are a couple of properties of the operator $B$:

1. $B$ is a monotone operator, as follows from the definition of enforceability:

$$W \subset W' \Longrightarrow B(W) \subset B(W').$$

2. $B$ maps compact sets into compact sets. If $(\alpha^k, w^k)$ is a sequence that decomposes $v^k$, and $\lim_k (\alpha^k, w^k) = (\alpha, w)$, then $(\alpha, w)$ decomposes $v = \lim_k v^k$. Hence, if $W$ is compact (so that sequences $(\alpha^k, w^k) \in \times_i \triangle A_i \times W^Y$ have convergent subsequences), $B(W)$ is closed, and clearly bounded.

**Definition 4** *The set $W \subset \mathbb{R}^n$ is **self-generating** if*

$$W \subset B(W).$$

The interest in self-generating sets follows from the following:

**Theorem 1** *If $W \subset \mathbb{R}^n$ is self-generating and bounded, $B(W) \subset E_\delta$.*

**Proof:** Because $W$ is self-generating, each $v \in B(W)$ is decomposed by some pair $(\alpha_v, w_v) \in \times_i \triangle A_i \times W^Y$. For each $v \in B(W)$, define the strategy $\sigma$, parametrized by $v' \in B(W)$, that starts at the beginning of the game in state $v$, and after any history $h^t$, given the current state

$v' \in B(W)$, specifies $\sigma(h^t) = \alpha_{v'}$, and moves to state $v'' = w_{v'}(y)$ in the next period, as a function of the realized signal $y$. We first show that states truly correspond to payoffs, i.e. that the payoff from playing $\sigma$ is indeed $v$. By definition of $\sigma$,

$$\begin{aligned}
v &= (1-\delta)\, u(\sigma(\varnothing)) + \delta \sum_{y^0} \pi\left(y^0 \mid \sigma(\varnothing)\right) w_v\left(y^0\right) \\
&= (1-\delta)\, u(\sigma(\varnothing)) + \delta \sum_{y^0} \pi\left(y^0 \mid \sigma(\varnothing)\right) \left[(1-\delta)\, u\left(\sigma\left(y^0\right)\right) + \delta \sum_{y^1} \pi\left(y^1 \mid \sigma\left(y^0\right)\right) w_{w_v(y^0)}\left(y^1\right)\right] \\
&= \cdots \\
&= (1-\delta) \sum_{s=0}^{t-1} \delta^s \sum_{h^s} \mathbb{P}_\sigma[h^s]\, u(\sigma(h^s)) + \delta^t \sum_{h^t} \mathbb{P}_\sigma\left[h^t\right] w_v\left[h^t\right],
\end{aligned}$$

where $\mathbb{P}_\sigma[h^t]$ is the probability of $h^t$ under $\sigma$ , and $w_v[h^t]$ is the state that is obtained after public history $h^t$, starting from state $v$, given the strategy $\sigma$. Because $w_v[h^t] \in W$ and $W$ is bounded, it follows that

$$v = \lim_{t\to\infty} (1-\delta) \sum_{s=0}^{t-1} \delta^s \sum_{h^s} \mathbb{P}_\sigma[h^s]\, u(\sigma(h^s)) = (1-\delta)\mathbb{E}_\sigma\left[\sum_{t=0}^{\infty} \delta^t u_i(a^t)\right],$$

as was to be shown. Similarly, $w_v[h^t]$ is the continuation payoff under $\sigma$ given public history $h^t$. Optimality of $\sigma_i$ then follows from the one-shot deviation principle, given that $\alpha_{i,v'}$ is optimal against $\alpha_{-i,v'}$, given continuation payoffs $w_{v'}$. □

Continuation payoffs from a PPE are equilibrium payoffs themselves. Hence, all equilibrium payoffs must be decomposable on $E_\delta$, and so $E_\delta \subset B(E_\delta)$. Therefore, $E_\delta$ is self-generating, and by the previous theorem, it follows that $B(E_\delta) \subset E_\delta$. Hence:

**Corollary 2** *It holds that*

$$E_\delta = B(E_\delta).$$

Hence, $E_\delta$ is a fixed-point of $B$. Note that, from Theorem 1, every bounded fixed-point of $B$ must be a subset of $E_\delta$ (because it must be self-generating), and so $E_\delta$ is actually the largest bounded fixed-point of $E_\delta$.

Because $V$, the set of feasible payoffs, is compact, and decomposable payoffs with respect to $V$ must be feasible,

$$V^1 = B(V) \subset V,$$

and $V^1$ is compact. More generally, by monotonicity of $W$, the sequence

$$V^{k+1} = B\left(V^k\right),$$

with $V^0 = V$, is (weakly) decreasing and compact. Furthermore, because

$$E_\delta \subset V \Longrightarrow E_\delta = B\left(E_\delta\right) \subset B\left(V\right) = V^1,$$

the set $V^1$ (and similarly $V^k$) is non-empty, containing $E_\delta$. Let $V^\infty = \lim_k V^k$. Because $V^k$ is a decreasing sequence of compact sets, $V^\infty$ is compact.

**Lemma 1** *The set $V^\infty$ is self-generating.*

**Proof:** Fix $v \in V^\infty$. Then $v \in V^k$ for all $k$, and so there exists a sequence $\left(\alpha^k, w^k\right)$ that decomposes $v$ with $w^k \in V^{k-1}$. We want to show that $v$ is decomposable on $V^\infty$, and the obvious candidate for decomposition is $(\alpha, w) = \lim_k \left(\alpha^k, w^k\right)$ (the sets are compact, so that we can assume that the sequence converges). We must show that $w \in V^\infty$. Suppose $w\left(y\right) \notin V^\infty$ for some $y$. Because $V^\infty$ is closed, there exists a compact neighborhood $\mathcal{N}$ of $w\left(y\right)$ such that $\mathcal{N} \cap V^\infty = \varnothing$. Because $w^k \to w$, there exists $K \in \mathbb{N}$ such that $w^k\left(y\right) \in \mathcal{N}$ for all $k \geq K$, and so, for all $k > K$,

$$\mathcal{N} \cap \left(\cap_{k' \leq k} V^{k'}\right) \neq \varnothing.$$ [8]

The collection $\left\{\mathcal{N}, V^k : k \in \mathbb{N}\right\}$ has the finite intersection property, so $\mathcal{N} \cap V^\infty \neq \varnothing$, a contradiction. □

It follows immediately that:

**Corollary 3** *It holds that*

$$E_\delta = V^\infty,$$

*and $E_\delta$ is compact.*

This suggests an immediate numerical algorithm for computing $E_\delta$ (see Judd, Yeltekin and Conklin, 2003). But comparative statics follow as well from this characterization, such as:

[8] Because $w^{k+1}(y) \in V^k = \cap_{k' \leq k} V^{k'}$.

**Lemma 2** *If $W \subset \mathbb{R}^n$ is bounded, convex, and self-generating for $\delta$, then it is also self-generating for $\delta' > \delta$.*

**Proof:** Suppose that $v \in W \subset B_\delta(W)$ is decomposed by $(\alpha, w)$ given $\delta$, and define

$$w' := \frac{\delta' - \delta}{\delta'(1-\delta)} v + \frac{\delta(1-\delta')}{\delta'(1-\delta)} w.$$

Note that $w'$ is in $W$, as $v, w \in W$ and $W$ is convex. Then

$$\begin{aligned} & (1-\delta') u(\cdot) + \delta' \sum_y \pi(y \mid \cdot) w'(y) \\ = \quad & \frac{\delta' - \delta}{1-\delta} v + \frac{1-\delta'}{1-\delta} \left[ (1-\delta) u(\cdot) + \sum_y \delta \pi(y \mid \cdot) w(y) \right], \end{aligned}$$

and so $\alpha$ is enforced by $w'$ given $\delta'$. □

Therefore, if $E_\delta$ is convex, then it is contained in $E_{\delta'}$ for all $\delta' > \delta$. But there are well-known examples in which $E_\delta$ is not convex, no matter how large $\delta$ is (Yamamoto, 2010). However, this is trivially the case if we assume that players have access to a public randomization device. In fact, we can then assume that $w$ only take as values the extreme points of $E_\delta$, and by Carathéodory's theorem, we only need to randomize over $n+1$ extreme points. Alternatively, APS show that attention can be restricted to extreme points if the set of signals is "large:"suppose that signals are distributed absolutely continuously with respect to Lebesgue measure on a subset of $\mathbb{R}^k$, for some $k \geq 1$. Then if $W \subset \mathbb{R}^n$ is compact, convex and self-generating, we can choose the decomposition $(\alpha, w)$ such that $w$ only takes values on the extreme points of $W$. While their proof relies on Lyapunov's theorem, a more transparent (but closely related) argument relies on Dubins-Spanier's "fair cake-cutting" theorem:

**Theorem 4** *Let $\mu_1, \ldots, \mu_n$ be nonatomic probability measures on a measurable space $(S, \Sigma)$. Given any $\beta_1, \ldots, \beta_m \geq 0$ with $\sum \beta = 1$, there is a partition $\{E_1, \ldots . E_m\}$ of $S$ such that for all $i = 1, \ldots, n$, and all $j = 1, \ldots, m$, $\mu_i(E_j) = \beta_j$. In fact, there is a sub-$\sigma$-algebra $\hat{\Sigma}$ on which all the measures agree, which is rich in the sense that for every $r \in [0,1]$, there is $E \in \hat{\Sigma}$, $\mu_i(E) = r$.*

This means that we can define a random variable that is uniformly distributed, and whose distribution is independent of the action profile chosen (pick $\mu_j = \pi(\cdot|a)$, where $j$ runs over the set of action profiles). Hence, the result follows. A more surprising result of APS is that, under some additional conditions, the continuation payoff *must* take values in the extreme points of $W$.

# III Characterization as $\delta \to 1$

We now strive to obtain a simpler characterization as $\delta \to 1$. By simpler, we mean: (i) a characterization that does not depend on $\delta$ (as we are taking limits, we may hope to be able to do so); (ii) a characterization that is not a fixed-point characterization.

We first define a local version of self-generation (first introduced by Fudenberg, Levine and Maskin, 1994) that simplify some of the analysis. We now index the operator $B$ by the relevant discount factor.

**Definition 5** *The set $W \subset \mathbb{R}^n$ is* ***locally self-generating*** *if $\forall v \in W, \exists \delta < 1$, and an open neighborhood $\mathcal{N}_v$ of $v$ such that $\mathcal{N}_v \cap W \subset B_\delta(W)$.*

Note that, for compact sets, local self-generation suffices to establish self-generation for some high enough discount factor, as we can take a finite subcover of the open cover $\mathcal{N}_v$, and use as discount factor the highest one for this finite subcover.

To eliminate the discount factor, let us proceed heuristically for now. If $\alpha$ is a Nash equilibrium of the one-shot game $\Gamma_\delta(w)$ with payoff $v \in E_\delta$, then, subtracting $\delta v$ on both sides of (3) and dividing through by $1-\delta$, we obtain

$$v = u(\alpha) + \sum_{y \in Y} \pi(y \mid \alpha) x(y),$$

where, for all $y$,

$$x(y) := \frac{\delta}{1-\delta}(w(y) - v), \text{ or } w(y) = v + \frac{1-\delta}{\delta} x(y).$$

Thus, provided that the equilibrium payoff set is convex, the payoff $v$ is also in $E_{\tilde{\delta}}$ for all $\tilde{\delta} > \delta$, because we can use as continuation payoff vectors $\tilde{w}(y)$ the re-scaled vectors $w(y)$ (see Figure 1). Conversely, provided that the normal vector to the boundary of $E_\delta$ varies continuously with the boundary point, then any set of payoff vectors $w(y)$ that lie in one of the half-spaces defined by this normal vector (*i.e.*, such that $\lambda \cdot (w(y) - v) \leq 0$, or equivalently, $\lambda \cdot x(y) \leq 0$) must also lie in $E_\delta$ for discount factors close enough to one. In particular, if we seek to identify the payoff $v$

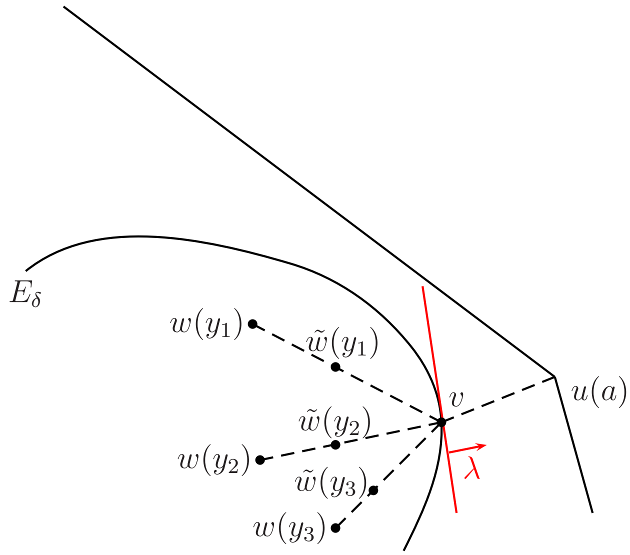


Figure 1: Continuation payoffs as a function of the discount factor

that maximizes $\lambda \cdot v$ on $E_\delta$ for $\delta$ close enough to 1, given $\lambda \in \mathbb{R}^n$, it suffices to compute the score

$$k(\lambda) := \sup_{x,v,\alpha} \lambda \cdot v,$$

such that $\alpha$ is a Nash equilibrium with payoff $v$ of the game $\Gamma(x)$ whose action sets are $A_i$ and payoff function is given by $u(a) + \sum_{y \in Y} \pi(y \mid a)x(y)$, and subject to the constraints $\lambda \cdot x(y) \leq 0$ for all $y$. Note that the discount factor no longer appears in this program. Furthermore, the unknown set $E_\delta$ no longer appears in the constraints.

We thus obtain a half-space $\mathcal{H}(\lambda) := \{v \in \mathbb{R}^n : \lambda \cdot v \leq k(\lambda)\}$ that contains $E_\delta$ for every $\delta$. This must be true for all vectors $\lambda \in \mathbb{R}^n$. Let $\mathcal{H} := \bigcap_{\lambda \in \mathbb{R}^n} \mathcal{H}(\lambda)$. We thus have:

$$\overline{\lim_{\delta \to 1}} \, E_\delta \subset \mathcal{H}.$$

What is more remarkable is that the converse inclusion holds as well, if $\mathcal{H}$ has non-empty interior. That is:

**Theorem 5** *If int $\mathcal{H} \neq \varnothing$, then*

$$\lim_{\delta \to 1} E_\delta = \mathcal{H}.$$

**Proof:** (Sketch of) The set $\mathcal{H}$ is compact and convex, and so it can be approximated by a convex set $W \subset \text{int } \mathcal{H}$ whose normal vector to the boundary of $W$ varies continuously with the boundary point. We show that $W$ is locally self-generating. If $v \in \text{int } W$, then there exists an open neighborhood $\mathcal{N}$ of $v$, $\mathcal{N} \subset \text{int } W$, and $\delta < 1$ such that $\forall v' \in \mathcal{N}$, $v' = (1-\delta)\, u(\alpha) + \delta w$ with $\alpha$ a Nash equilibrium of the stage game and $w \in W$.

Assume now that $v$ is a boundary point. Let $\lambda \in \mathbb{R}^n$ denote the normal vector, and let $k := \lambda \cdot v < k(\lambda)$ (recall that $W \subset \text{int } \mathcal{H}$). Fix $(\alpha, w^*)$ that decomposes some $v^*$ such that $\lambda \cdot v^* \in (k, k(\lambda))$. Note that $\lambda \cdot x^*(y) \leq 0$ (where $x^* := \delta(w^* - v^*) / (1-\delta)$). To make those strict inequalities, pick $v'$ such that $k < \lambda \cdot v' < \lambda \cdot v^*$ and note that $(\alpha, w')$ decomposes $v'$, where

$$w' := v' + \frac{1-\delta}{\delta} x', \quad x' := x^* - v^* + v'.$$

Furthermore, $\lambda \cdot x' \leq \lambda \cdot (v' - v^*) =: -\kappa < 0$. We must find $\varepsilon > 0$, $\delta < 1$ such that, for all $v'' \in W$, $\|v'' - v\| < \varepsilon$, $v''$ is decomposed by $(\alpha, w'')$, where

$$w'' = v'' + \frac{1-\delta}{\delta}(x' - v' + v''),$$

with $w'' \in W$. The difficult constraint is $w'' \in W$. Note that $\max_y \|w''(y) - v''\| = O(1-\delta)$, while

$$\lambda \cdot w''(y) \leq \lambda \cdot v'' - \frac{1-\delta}{\delta}\kappa.^{9}$$

The existence of such $\delta$, $\varepsilon$ then follows from the smoothness of $W$ at $v$.[10] See Figure 2. $\square$

Note that this result does not only characterize the limit, it establishes the existence of this limit.

This characterization can be adapted to the case in which some of the players are "short-run," *i.e.*, myopic ($\delta = 0$). Suppose that players $i = 1, \ldots, L$, $L \leq n$, are long-run players, whose objective is to maximize the average discounted sum of rewards, with discount factor $\delta < 1$. Players $j \in SR := \{L+1, \ldots, n\}$ are short-run players, each representative of which plays only once. Let

$$B : \times_{i=1}^{n} \triangle A_i \to \times_{j=L+1}^{n} \triangle A_j$$

[9] To see this, note that, since $v'' \in W$, $\lambda \cdot v'' \leq k < \lambda v'$.

[10] All that is required is that the boundary of $W$ has no "kink" at $v$, as what we need is that points whose scores are at least $\frac{1-\delta}{\delta}\kappa$ lower than the score of a point in $W$, yet within a distance of order $1-\delta$ of this point, be also in $W$.

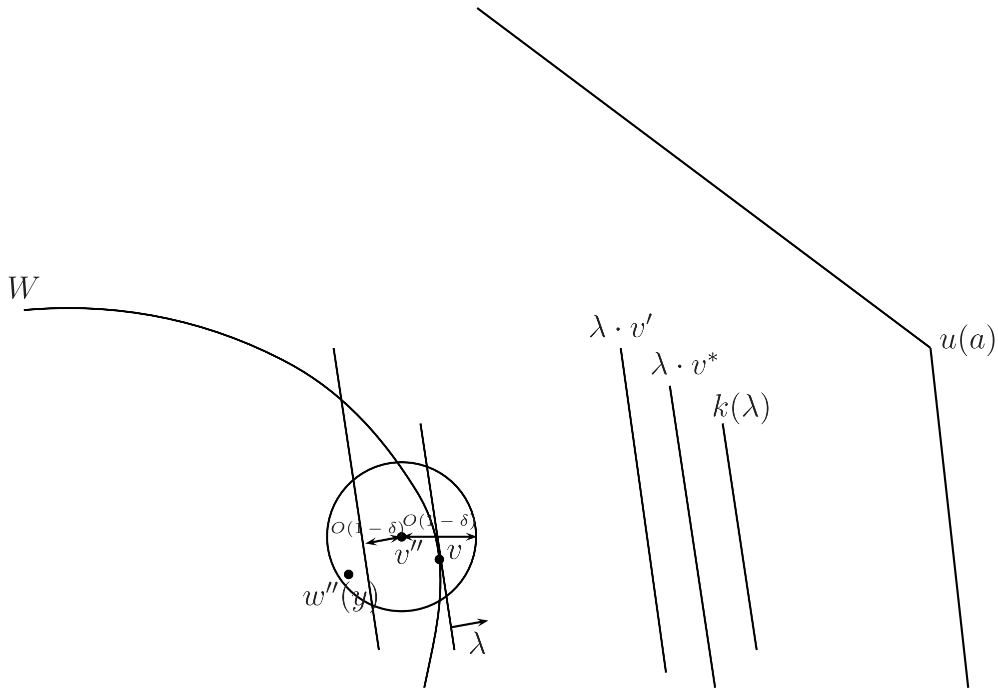


Figure 2: Construction in the sketch

be the correspondence that maps any mixed action profile $(\alpha^1, \ldots, \alpha^n)$ for the long-run players to the corresponding static equilibria for the short-run players in state. That is, for each $\alpha \in \text{graph}B$, and each $j > L$, $\alpha^j$ maximizes $u_j(\cdot, \alpha_{-j})$. The characterization goes through if we "ignore" the short-run players and simply require that $v$ be a Nash equilibrium payoff of the game $\Gamma(x)$ for the long-run players, achieved by some $\alpha \in$ graph $B$.

# IV The Folk Theorem

Given our characterization of the limiting payoff set, we are left with the task of finding sufficient conditions under which $\mathcal{H}$ is equal to the feasible and individually rational payoff set. Recall that the main ingredient is the following maximization program $\mathcal{P}(\lambda)$, parameterized by $\lambda \in \mathbb{R}^n$:

$$\sup_{v,x,\alpha} \lambda \cdot v,$$

where the supremum is taken over all $v \in \mathbb{R}^I$, $x : Y \to \mathbb{R}^n$, and $\alpha \in \times_i \triangle A_i$ such that

**(i)** $\alpha$ is a Nash equilibrium with payoff $v$ of the game $\Gamma(x)$;

**(ii)** For each $y \in Y$, $\lambda \cdot x(y) \leq 0$.

For a fixed $\alpha$, the feasible set of $\mathcal{P}(\lambda)$ is non-empty if and only if $\alpha$ is **admissible**, in the sense that, for all $i$, if there exists $\nu \in \triangle A_i$ such that, for all $y$,

$$\sum_k \nu^k \pi(y \mid a_i^k, \alpha_{-i}) = \pi(y \mid \alpha),$$

(here, $k$ runs over $i$'s actions) then

$$\sum_k \nu^k u_i(a_i^k, \alpha_{-i}) \leq u_i(\alpha).$$

Indeed, it follows from Fan (1956) that there exists $x : Y \to \mathbb{R}^n$ such that $\alpha$ is a Nash equilibrium of the game $\Gamma(x)$ if and only if $\alpha$ is admissible. Considering any $i$ for which $\lambda_i \neq 0$, we may add (or subtract if $\lambda_i < 0$) a constant to $x_i(y)$, independent of $y$, so that the constraint **(ii)** is satisfied.

It follows from duality that $\mathcal{P}(\lambda)$ is equivalent to (*i.e.*, gives the same value as) $\tilde{\mathcal{P}}(\lambda)$ given by

$$\sup_{\alpha \in \times_i \triangle A_i, \alpha \text{ admissible}} \min \sum_i \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right),$$

where the minimum is over $(\hat{\alpha}_i)_{\{i:\lambda_i \neq 0\}}$, $\sum_k \hat{\alpha}_i\left(a_i^k\right) = 1$, with $\lambda_i > 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 1)$, $\lambda_i < 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 1)$, and

$$\lambda_i \neq 0 \Rightarrow \hat{\pi}\left(y\right) := \pi\left(y \mid \hat{\alpha}_i, \alpha_{-i}\right) \geq 0.$$

**Proof:** (of the dual representation)
Fix throughout some strategy $(\alpha)$ such that $\alpha$ is admissible. We can rewrite $\mathcal{P}(\lambda)$ as

$$\max_{x,v} \lambda \cdot v$$

over $x$ and $v$ such that, for all $i$,

$$\sum_y \pi\left(y \mid \alpha\right) x_i\left(y\right) - v_i = -u_i\left(\alpha\right),$$

and, for all $i, k$,

$$\sum_y \left[\pi\left(y \mid a_i^k, \alpha_{-i}\right) - \pi\left(y \mid \alpha\right)\right] x_i\left(y\right) \leq u_i\left(\alpha\right) - u_i\left(a_i^k, \alpha_{-i}\right),$$

as well as, for all $y$,

$$\lambda \cdot x(y) \leq 0.$$

This is a linear program for $(x, v)$. The first set of constraints ensure that $\alpha$ yields the payoff $v$, the second that playing $\alpha$ is a Nash equilibrium, and the third is the same constraint as **(ii)**. Because we assumed that $\alpha$ is admissible, the feasible set is non-empty, and because the value of this program is bounded above by $k(\lambda)$, it is finite. We shall consider the dual of this linear program. It is

$$\min -\sum_i \gamma_i u_i(\alpha) + \sum_{i,k} \nu_i^k \left(u_i(\alpha) - u_i\left(a_i^k, \alpha_{-i}\right)\right)$$

over $\gamma_i \in \mathbb{R}, \nu_i^k \geq 0, \eta_y \geq 0$, such that, for all $i, y$,

$$\pi(y \mid \alpha)\gamma_i - \sum_k \left[\pi(y \mid \alpha) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right]\nu_i^k + \lambda_i \eta_y = 0,$$

where $k$ runs through the actions of player $i$, and

$$-\gamma_i = \lambda_i.$$

Let $B := \{i : \lambda_i \neq 0, i \in I\}$ and $B' := \{i : \lambda_i = 0, i \in I\}$. Substituting $-\gamma_i = \lambda_i$ into the dual program, we get

$$\min \sum_{i \in B}\left[\lambda_i u_i(\alpha) + \sum_k \left(u_i(\alpha) - u_i\left(a_i^k, \alpha_{-i}\right)\right)\nu_i^k\right] + \sum_{i \in B'}\left[\sum_k \left(u_i(\alpha) - u_i\left(a_i^k, \alpha_{-i}\right)\right)\nu_i^k\right]$$

over $\nu_i^k \geq 0, \eta_y \geq 0$, such that, for all $y$ and $i \in B$,

$$\pi(y \mid \alpha) + \sum_k \left[\pi(y \mid \alpha) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right]\frac{\nu_i^k}{\lambda_i} = \eta_y,$$

for all $y$ and $i \in B$,

$$\sum_k \left[\pi(y \mid \alpha) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right]\nu_i^k = 0.$$

For $i \in B'$, $\sum_k \left[\pi(y \mid \alpha) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right]\nu_i^k = 0$ can be satisfied by setting

$$\nu_i^k = \alpha_i(a_i^k) \geq 0.$$

Moreover, by admissibility,

$$\sum_k \left(u_i\left(\alpha\right) - u_i\left(a_i^k, \alpha_{-i}\right)\right)\nu_i^k \geq 0$$

whenever $\sum_k \left[\pi\left(y \mid \alpha\right) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right]\nu_i^k = 0$. Hence, we can remove the constraints associated with $i \in B'$.

For $i \in B$, define $\xi_i^k := \nu_i^k / \lambda_i$. The dual program can be reduced to

$$\min \sum_{i \in B} \lambda_i \left[u_i\left(\alpha\right) + \sum_k \left(u_i\left(\alpha\right) - u_i\left(a_i^k, \alpha_{-i}\right)\right)\xi_i^k\right]$$

over $\nu_i^k \geq 0, \eta_y \geq 0$, such that, for all $y$ and $i \in B$

$$\pi\left(y \mid \alpha\right) + \sum_k \left(\pi\left(y \mid \alpha\right) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right)\xi_i^k = \eta_y.$$

Define $\hat{\alpha}_i \in \mathbb{R}^{|A_i|}$ by, for all $a_i^k$, $i \in B$,

$$\hat{\alpha}_i\left(a_i^k\right) = \alpha_i\left(a_i^k\right) + \alpha_i\left(a_i^k\right)\sum_{k'} \xi_i^{k'} - \xi_i^k.$$

It can be easily verified that

$$\pi\left(y \mid \hat{\alpha}_i, \alpha_{-i}\right) = \pi\left(y \mid \alpha\right) + \sum_k \left(\pi\left(y \mid \alpha\right) - \pi\left(y \mid a_i^k, \alpha_{-i}\right)\right)\xi_i^k.$$

Note that $\sum_k \hat{\alpha}_i\left(a_i^k\right) = 1$ for all $i \in B$. We can rewrite our problem as

$$\min \sum_{i \in B} \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right),$$

over $\left(\hat{\alpha}_i\right)_i$, $\sum_k \hat{\alpha}_i\left(a_i^k\right) = 1$, with $\lambda_i > 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 1)$, $\lambda_i < 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 1)$, as well as, and $\eta_y \geq 0$, such that, for all $y$ and $i \in B$,

$$\pi\left(y \mid \hat{\alpha}_i, \alpha_{-i}\right) = \eta_y. \tag{4}$$

Since $\sum_{i \in B} \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right) = \sum_{i \in B} \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right) + \sum_{i \in B'} \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right)$ no matter how we define $u_i(\hat{\alpha}_i, \alpha_{-i})$ for $i \in B$, adding back $0 = \sum_{i \in B'} \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right)$ we can rewrite our problem without using $\eta_y$ as follows:

$$\min \sum_i \lambda_i u_i\left(\hat{\alpha}_i, \alpha_{-i}\right),$$

over $(\hat{\alpha}_i)_i$, $\sum_k \hat{\alpha}_i\left(a_i^k\right) = 1$, with $\lambda_i > 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 1)$, $\lambda_i < 0 \Rightarrow (\alpha_i(a_i^k) = 0 \Rightarrow \hat{\alpha}_i(a_i^k) \geq 0, \alpha_i(a_i^k) = 1 \Rightarrow \hat{\alpha}_i(a_i^k) \leq 1)$, and

$$\lambda_i \neq 0 \Rightarrow \hat{\pi}(y) := \pi\left(y \mid \hat{\alpha}_i, \alpha_{-i}\right) \geq 0$$

Taking the supremum over admissible $(\alpha_s)_s$, this gives us precisely $\tilde{\mathcal{P}}(\lambda)$. □

In a slight departure from the notations of the set of notes on perfect monitoring, let $\underline{V}$ denote the set of feasible and weakly individually rational payoff vectors. Denote by $e_i$ the $i$-th basis vector in $\mathbb{R}^n$. The direction $\lambda$ is a **coordinate direction** if $\lambda = \lambda_i e_i$ for some $\lambda_i \in \mathbb{R}$, $\lambda_i \neq 0$. It is a **non-coordinate direction** otherwise. We denote the set of non-coordinate directions by $\Lambda^{nc}$, and of coordinate directions by $\Lambda^c$. Denote such a direction $\lambda^i$. Finally, let $Ex(A)$ denote the set of (necessarily pure) action profiles achieving some extreme point of the feasible payoff set $V$.

Under what conditions does $\lim_{\delta \to 1} E_\delta = \underline{V}$? Assuming throughout that $\underline{V}$ has non-empty interior, it reduces to finding conditions under which, in every direction $\lambda \in \mathbb{R}^n$,

$$k(\lambda) = \max_{v \in \underline{V}} \lambda \cdot v. \tag{5}$$

Depending on the direction $\lambda$, the maximum on the right-hand side is achieved either by some $a \in Ex(A)$, or some minmax action profile $\underline{\alpha}^i$ (for coordinate directions $\lambda^i = \lambda_i e_i$, $\lambda_i < 0$).

Therefore, among the set of sufficient conditions for a folk theorem, we start with:

**Assumption A1**: For every $i$, some $\underline{\alpha}^i$, is admissible.

Let us define the matrix $\Pi_i\left(\alpha_{-i}\right)$ as the $|A_i| \times |Y|$-matrix whose $(a_i, y)$-th entry is $\pi\left(y \mid a_i, \alpha_{-i}\right)$. Further, given $\alpha$, the matrix $\Pi_{ij}\left(\alpha\right)$ is defined as

$$\Pi_{ij}\left(\alpha\right) = \begin{pmatrix} \Pi_i\left(\alpha_{-i}\right) \\ \Pi_j\left(\alpha_{-j}\right) \end{pmatrix}$$

Note that this matrix has maximal rank $|A_i| + |A_j| - 1$, because $\alpha_i \Pi_i\left(\alpha_{-i}\right) = \alpha_j \Pi_j\left(\alpha_{-j}\right)$.

Admissibility for some $\alpha \in \times_i \triangle A_i$ is automatically satisfied if the matrix $\Pi_i\left(\alpha_{-i}\right)$ has full row rank for every $i$: this is what Fudenberg, Levine and Maskin define as **individual full rank** for $\alpha$. But admissibility is clearly a weaker requirement.

If $\lambda$ is a coordinate direction $\lambda^i = \lambda_i e_i$, $\lambda_i < 0$, no further assumptions are necessary for (5), since player $i$ is the only one whose action the minimum is taken with respect to in the dual, but then again, $\underline{\alpha}^i$ dictates that he takes a best-reply (which is what the minimum calls for, given that $\lambda_i < 0$).

How about other directions? We need to make sure that the only $\hat{\alpha}_i$ allowed by the constraints in the dual is actually $\alpha_i$ (we are interested, of course, in $\alpha = a \in Ex(A)$).

Let $Q_i(a) := \{\pi(\cdot \mid a_{-i}, a_i^k) \mid a_i^k \in A_i \setminus \{a_i\}\}$ be the set of distributions over signals as player $i$'s action varies over all his actions but $a_i$. Let $C_i(a)$ denote the convex cone with vertex 0 spanned by $Q_i(a) - \pi(\cdot \mid a)$. Note now that the restriction on $\hat{\alpha}$, when $\alpha = a$ is pure, is that $\pi(\cdot \mid \hat{\alpha}_i, a_{-i}) - \pi(\cdot \mid a) \in -C_i(a)$ whenever $\lambda_i > 0$, and $\pi(\cdot \mid \hat{\alpha}_i, a_{-i}) - \pi(\cdot \mid a) \in C_i(a)$ whenever $\lambda_i < 0$. Finally, we need admissibility, for which it suffices that 0 is not a conical combination of $Q_i(a) - \pi(\cdot \mid a)$, *i.e.*, there exists no positive weights on actions of $i$ different that $a_i$ such that the resulting average distribution over signals coincide with the one induced by $a_i$, given $a_{-i}$. We thus impose:

**Assumption A2**: For every $a \in Ex(A)$, every pair of players $i, j$, $C_i(a) \cap -C_j(a) = C_i(a) \cap C_j(a) = \{0\}$, and 0 is not a conical combination of $Q_i(a) - \pi(\cdot \mid a)$.

The next theorem is now an immediate corollary.

**Theorem 6** *(The folk theorem) If (i) $\underline{V}$ has non-empty interior, and (ii) Assumptions **A1** and **A2** hold, then*

$$\lim_{\delta \to 1} E_\delta = \underline{V}.$$

This theorem was proved under slightly stronger assumptions by Fudenberg, Levine and Maskin (1994), and under the currently stated ones by Kandori and Matsushima (1998). The dual representation suggests further possible weakenings, but we shall not do so. On the contrary, we note that, from the dual representation, a stronger assumption can be made that implies **A2**, namely, we may assume that the profile $a$ has **pairwise full rank**: that is, for $i$ and $j$ the matrix $\Pi_{ij}(a)$ has rank $|A_i| + |A_j| - 1$ for all pairs $i, j$. This is the assumption originally made by Fudenberg, Levine and Maskin (1994).

We may now finally come back to our original example. How come the folk theorem failed? Leaving aside the fact that the characterization of the asymptotic payoff set does not apply in this case ($\mathcal{H}$ has empty interior, as the set of equilibrium payoffs is an interval) the problem

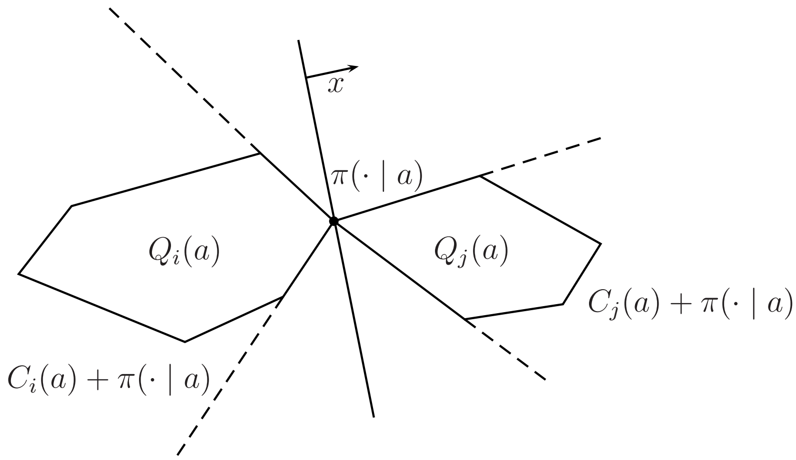


Figure 3: On the role of $C_i(a) \cap C_j(a) = \{0\}$

is that one of the rank conditions failed. In fact, both rank conditions failed: there are only two signals, but a continuum of actions, so clearly players cannot be made indifferent across all actions. But individual full rank is not the main problem, as deviations from full effort can still be statistically detected. The main problem is that players cannot identify deviators (which pairwise full rank requires), and therefore, cannot provide incentives without punishing everyone; the key insight behind FLM is that providing incentives does not require losing any efficiency, if as a function of the signal, the "aggregate payoff" is re-distributed across players so as to keep the sum fixed. This was, as mentioned, the key idea behind pairwise directions and pairwise full rank.

Figures 3 and 4 illustrate the role of the assumptions. These figures represent probability distributions and sets of such distributions. As is clear from Figure 3, if $C_i(a) \cap C_j(a) = \{0\}$, one can find a direction $x$ such that, by punishing one player in that direction while simultaneously rewarding his opponent, incentives are aligned: to avoid the punishment (or reap the reward) both players have an incentive to select $a$ in terms of the actions available to them, in terms of the signal distributions they can generate via unilateral deviations. Similarly, Figure 4 illustrates that, if $C_i(a) \cap -C_j(a) = \{0\}$, one can find a direction $x$ such that incentives are aligned when both players' payments go in the same direction, so that they are both simultaneously punished or rewarded.

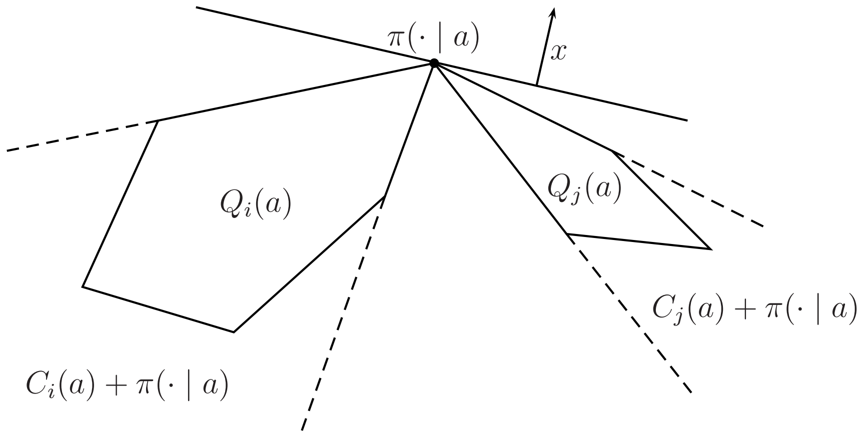


Figure 4: On the role of $C_i(a) \cap -C_j(a) = \{0\}$

# V Some Qualifications and Extensions

## A What if $\operatorname{int} \mathcal{H} = \varnothing$?

This may happen. Fortunately, the scoring algorithm by Fudenberg and Levine (1994) has been refined by Fudenberg, Levine and Takahashi (2007) to cover this case. The extension is sufficiently intuitive that no proof will be given. Given a set $X \subseteq \mathbb{R}^n$, let $\operatorname{aff} X := \{\sum_{k=1}^{K} \lambda_k x_k : K \in \mathbb{N}, x_k \in X, \sum_{k=1}^{K} \lambda_k = 1\}$ denote the affine hull of $X$. (Note that $\lambda_k$ can take negative values.) That is, an affine hull is simply the sum of some vector $y \in \mathbb{R}^n$ and a linear subspace $U(X) \subseteq \mathbb{R}^n$, namely $\operatorname{aff} X = y + U(X)$. The dimension of $\operatorname{aff} X$ is defined to be the dimension of $U(X)$. Given some affine hull $X$, let $\mathcal{P}(\lambda, X)$ be the program

$$k(\lambda, X) := \sup_{x,v,\alpha} \lambda \cdot v,$$

such that $\alpha$ is a Nash equilibrium with payoff $v$ of the game $\Gamma(x)$ whose action sets are $A_i$ and payoff function is given by $u(a) + \sum_{y \in Y} \pi(y \mid a) x(y)$, and subject to the constraints $\lambda \cdot x(y) \leq 0$ for all $y$, as well as $x(y) \in X$ for all $y$.

This is a simple modification of the program that we have defined in Section III, with the added constraint that $x$ takes values in $X$. Let $\mathcal{H}(\lambda, X) := \{v \in \mathbb{R}^n : \lambda \cdot v \leq k(\lambda)\}$ and $\mathcal{H}(X) := \bigcap_{\lambda \in \mathbb{R}^n, \lambda \in U(X)} \mathcal{H}(\lambda, X) \cap X$.

We now define the finite sequence $\{X^k\}_{k=1}^{K}$ as follows: $X^1 = \mathbb{R}^n$; for $k \geq 1$, if $\dim \mathcal{H}(X^k) =$

|  | $C$ | $D$ |
|---|---|---|
| $C$ | $2, 2$ | $-1, 3$ |
| $D$ | $3, -1$ | $0, 0$ |

Figure 5: Prisoner's dilemma

$\dim X^k$, set $X^{k+1} = \mathcal{H}(X^k)$ and $K = k$ (so, stop the sequence); if $\dim\mathcal{H}(X^k) < \dim X^k$, set $X^{k+1} = \text{aff}\mathcal{H}(X^k)$ and continue.

Because at each step of the sequence (as long as it does not terminate) the dimension of $X^k$ diminishes by at least one, the sequence cannot be of length larger than $n$. It turns out that

$$\lim_{\delta \to 1} E_\delta = \mathcal{H}(X^K).$$

We note that Theorem 5 is the special case in which the procedure ends in one step, because $\text{int}\mathcal{H} = \text{int}\mathcal{H}(\mathbb{R}^n) \neq \varnothing$. As a corollary of this generalization, it follows that $\lim_{\delta \to 1} E_\delta$ always exists.

# B Are the Full Rank Assumptions Necessary?

The example of the first section and the ensuing analysis strongly suggests that some conditions such as individual full rank and pairwise full rank might not just be sufficient, but also necessary to support cooperation when incentives in the stage game are misaligned. Yet the restriction to public strategies raises the possibility that these assumptions are driven by the solution concept (PPE) rather than the "fundamentals," and that perhaps these assumptions could be dispensed with a weaker solution concept.

By now, there are some nice examples of games in which private strategies improve on public ones (See, for instance, Kandori and Obara, 2006). Indeed, it is "easy" to construct examples in which cooperation can be supported in the two-player prisoner's dilemma example despite lack of pairwise full rank.[11] Consider the game in Figure 5. Suppose that there are two signals, $y = \underline{y}, \bar{y}$, with the probability of the signal $\bar{y}$ being $p$ if both agents cooperate, $q$ if only one does, and $r$ when neither cooperates. Assume that $1 > p > q > r > 0$, so that the likelihood of $\bar{y}$ increases with the number of cooperators. We assume sufficient noise; specifically, suppose that $2(p - q) < 1 - p$ and $r > \max\{2q - p, (3p^2 + 4q^2 - 6pq)/p\}$.

[11] This possibility contradicts the famous example of Radner, Myerson and Maskin (1986), but their analysis is restricted to PPE, seemingly unwittingly.

Solving for the limit set (as discounting vanishes) of public perfect equilibrium payoffs – equilibria in public strategies– has become a straightforward exercise, thanks to the characterization of Fudenberg and Levine (1994) and its refinement by Fudenberg, Levine and Takahashi (FLT, 2004). Consider the score in direction $(-1, 1)$, namely the maximum of

$$u^2(\alpha) - u^1(\alpha),$$

where $\alpha = (\alpha^1, \alpha^2)$ is the strategy profile played (*i.e.*, $\alpha^i$ is the probability that player $i$ plays $C$). It is then a simple matter of algebra to check that, given our restrictions on the parameters, maximizing this difference over action profiles $\alpha$, subject to incentive compatibility and budget-balance yields a score of 0, achieved by setting $\alpha = x = 0$. Fudenberg and Levine's result implies that all payoff vectors $(v^1, v^2)$ must then satisfy $v^2 - v^1 \leq 0$. By considering the direction $(-1, 1)$, we then get that it must hold that $v^1 = v^2$: all equilibrium payoffs must be symmetric payoffs. This implies that int $\mathcal{H} = \varnothing$, and Theorem 5 does not apply, although inspection of the proof shows that one direction of the theorem does, namely, $\limsup_{\delta \to 1} E_\delta \subset \mathcal{H}$. This is where the refinement of FLT applies. Loosely speaking, FLT's theorem states that, whenever int $\mathcal{H} = \varnothing$, one should recompute the scores $\mathcal{P}(\lambda)$ by adding one more constraint, namely, the vector $x$ lies in the smallest affine subspace of $\mathbb{R}^n$ that contains $\mathcal{H}$. Taking the intersection over resulting half spaces, we then obtain a "new" set $\mathcal{H}'$ contained in this affine subspace, and provided this one has non-empty interior, *relative* to the affine subspace, we have identified the limit equilibrium payoff set.[12]

Applying this refinement, we may thus re-visit the efficient direction $(1, 1)$, and maximize

$$u^1(\alpha) + u^2(\alpha),$$

subject to budget-balance, $x^1(y) + x^2(y) \leq 0$, and symmetry, $x^1(y) = x^2(y)$. Again, simple algebra show that, given the requirement of symmetry, we can do no better than set $\alpha = 0$ and $x = 0$: all equilibrium payoffs must satisfy $v^1 + v^2 \leq 0$, and we are done, as this implies that the unique PPE payoff is $(0, 0)$, independently of the discount factor. Yet at the cost of a more complicated construction, one can prove that cooperation can be approached, by relying on ideas inspired by private monitoring.

Yet *some* minimum assumptions are required. Even some version of pairwise full rank is

[12]If $\mathcal{H}'$ does not have non-empty relative interior, this procedure must be iterated, requiring now the constraint that $x$ must lie in the smallest affine subspace of $\mathbb{R}^n$ that contains $\mathcal{H}'$. In this fashion, the limit set of PPE can always be solved for.

| | $W$ | $S$ |
|---|---|---|
| $W$ | $w-c, w-c, B$ | $w-c, w, 0$ |
| $S$ | $w, w-c, 0$ | $w, w, 0$ |

Table 1: Open

| | $W$ | $S$ |
|---|---|---|
| $W$ | $b, 0, 0$ | $b, 0, 0$ |
| $S$ | $b, 0, 0$ | $b, 0, 0$ |

Table 2: Delegate to $A_1$

| | $W$ | $S$ |
|---|---|---|
| $W$ | $0, b, 0$ | $0, b, 0$ |
| $S$ | $0, b, 0$ | $0, b, 0$ |

Table 3: Delegate to $A_2$

required, at least if there are three players or more. The following is an example due to Tomala (unpublished). Consider the following matrix game.

The interpretation is as follows. It is a game between a principal ($P$) and two agents ($A_1$, $A_2$). The principal is the owner of a firm and the agents are employees. Each day, the principal decides whether to open the firm or not. If he does, agents may work or shirk, these choices being unobservable. The firm produces an output only when both agents work. An agent who shirks gets his daily wage without paying the effort cost. If the principal decides not to open the firm, he has only two options, delegating the firm to agent $A_1$ or to agent $A_2$.

There, $B$ is the expected profit of the principal when he opens the firm and both player work, $w$ is the wage paid to each agent and $c$ is the cost of effort. We assume $w - c > 0$. Then, $b$ is the expected profit of either agent when he operates the firm by himself. We assume that $(Open, W, W)$ is the only surplus-efficient outcome, that is, $B + 2w - 2c > 2w$ (or $B > 2c$) and $B > b$. Two important properties of this one-shot game are:

1. The minmax is 0 for each player, so that all feasible payoffs are individually rational.
2. The set of feasible payoffs has non-empty interior.

The question now is whether the efficient outcome can be (approximately) implemented by an equilibrium of the repeated game when players are patient. Our main point is that, if the principal cannot tell the difference between agent $A_1$ shirking or agent $A_2$ shirking, then no matter how he tries to punish, he rewards one of the agents.

Specifically, assume that the action of the principal and the output (*i.e.*, the principal's payoff) are publicly observed. The public signal is given by the following table. We see that, for any (possibly mixed) action profile $(a_1, a_2, a_P)$, the public signals are the same under $(S, a_2, a_P)$ and under $(a_1, S, a_P)$. It is neither essential that signals be public, nor be a deterministic function of actions.

Under this structure, any (not necessarily public perfect) equilibrium payoff $v = (v_1, v_2, v_P)$ satisfies $v_1 + v_2 \geq \min\{2w, b\}$. Let $\sigma$ be an equilibrium of the repeated game. For $i = 1, 2$, let $\tau_i$ be the deviation of agent $A_i$ that plays $S$ always. Under the signaling structure, the distribution

| | $W$ | $S$ |
|---|---|---|
| $W$ | $(O, B)$ | $(O, 0)$ |
| $S$ | $(O, 0)$ | $(O, 0)$ |

Table 4: Open

| | $W$ | $S$ |
|---|---|---|
| $W$ | $(D_1, 0)$ | $(D_1, 0)$ |
| $S$ | $(D_1, 0)$ | $(D_1, 0)$ |

Table 5: Delegate to $A_1$

| | $W$ | $S$ |
|---|---|---|
| $W$ | $(D_2, 0)$ | $(D_2, 0)$ |
| $S$ | $(D_2, 0)$ | $(D_2, 0)$ |

Table 6: Delegate to $A_2$

of signals of the principal is the same under $(\tau_1, \sigma_{-1})$ as under $(\tau_2, \sigma_{-2})$. Thus, the actions of the principal also have the same distributions under these two strategy profiles. For each action of the principal $a \in \{O, D_1, D_2\}$, let

$$f_\sigma(a) = \mathbb{E}_\sigma \left[ \sum_{t \geq 1} (1-\delta)\delta^{t-1} \mathbf{1}_{\{a_t = a\}} \right]$$

be the expected discounted average number of times where the action $a$ is played. Since the deviations $\tau_1, \tau_2$ are not identifiable, we have

$$\forall a \in \{O, D_1, D_2\}, f_{(\tau_1, \sigma_{-1})}(a) = f_{(\tau_2, \sigma_{-2})}(a) =: f(a).$$

Now, $U_i(\tau_i, \sigma_{-i}) = wf(O) + bf(D_i)$, and from the equilibrium condition $v_i \geq U_i(\tau_i, \sigma_{-i})$. It follows that

$$v_1 + v_2 \geq wf(O) + bf(D_1) + wf(O) + bf(D_2) \geq \min\{2w, b\}(f(O) + f(D_1) + f(D_2)) = \min\{2w, b\}.$$

As an immediate corollary, if $2w - 2c < b$, then the efficient outcome $(w - c, w - c, B)$ is not an equilibrium outcome of the repeated game, and equilibrium payoffs are bounded away from efficiency.

# C Rates of Convergence

One might conclude from the folk theorems under imperfect public monitoring that, under appropriate rank conditions, it is irrelevant whether monitoring is imperfect or not. While this is certainly a valid viewpoint in the limit as $\delta \to 1$, it must be qualified for a fixed discount factor, and in terms of convergence rates.

First, for a fixed discount factor, Kandori (1992) has shown that as monitoring improves in the sense of Blackwell, the set of PPE weakly expands: hence, better monitoring cannot hurt. This improvement is quantified in Hörner and Takahashi (2015) in terms of convergence rate. They show that the set of PPE payoffs $E_\delta$ approaches the set of individually rational payoff vectors $\underline{V}$ at rate (at least) $(1-\delta)^{1/2}$ under perfect monitoring, and that this rate is tight (namely, examples can be found where this is precisely the rate of convergence). On the other hand, under public monitoring, this rate dips to $(1-\delta)^{1/4}$: under standard individual and pairwise full rank assumptions, $E_\delta \to \underline{V}$ at least as fast as rate $(1-\delta)^{1/4}$,[13] and examples can be found where convergence occurs precisely at this rate. Hence, imperfect monitoring comes at a cost.

# VI Literature

The literature on repeated games with imperfect public monitoring was motivated, among others, by Green and Porter 1984 ("Noncooperative Collusion under Imperfect Price Information," *Econometrica*, **52**, 87–100). The initial example, which illustrates the difference between good news and bad news, is due to Abreu, Milgrom and Pearce 1991 ("Information and Timing in Repeated Partnerships," *Econometrica*, **59**, 1713–1733).

Abreu, Pearce and Stacchetti developed the main ideas behind self-generation (Abreu, D., D. Pearce, and E. Stacchetti 1990, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, **58**, 1041–1063), though similar ideas were introduced in Mertens, J.-F. and T. Parthasarathy 1987 ("Equilibria for Discounted Stochastic Games," C.O.R.E. Discussion Paper 8750). As mentioned, the operator was introduced by Shapley in 1953 ("Stochastic Games," *Proceedings of National Academy of Science*, **39**, 1095–1100).

The topological structure of $E_\delta$ is not well-studied. As mentioned in the text, $E_\delta$ need not be increasing in $\delta$ when one does not assume a public randomization device, no matter whether one restricts attention to arbitrarily high discount factors or not, and a counter-example under perfect monitoring can be found in Yamamoto (2010, "The Use of Public Randomization in Discounted Repeated Games," *International Journal of Game Theory*, **39**, 431–443).

The numerical algorithm mentioned in the text is developed in Judd, Yeltekin and Conklin, 2003 ("Computing Supergame Equilibria," *Econometrica*, **71**, 1239–1254). Finally, the "fair cake-cutting" algorithm that is mentioned in Section II is due to Dubins and Spanier 1961 ("How to cut a cake fairly," *American Mathematical Monthly*, **68**, 1–17). Dubins and Spanier prove the first part of Theorem 4, involving a finite partition. The second part of Theorem 4, involving an arbitrary number $r \in [0,1]$, is due to Border, Ghirardato and Segal 2008 ("Unanimous Subjective Probabilities," *Economic Theory*, **34**, 383–387).

[13]This means that $d(E_\delta, \underline{V}) \leq M(1-\delta)^{1/4}$, for some constant $M$ and all $\delta < 1$.

The limit characterization of equilibrium payoffs based on scores is due to Fudenberg and Levine 1994 ("Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, **62**, 103–135), who went on to use it, as we have done in these notes, to prove the folk theorem in Fudenberg, Levine and Maskin 1994 ("The Folk Theorem with Imperfect Public Information," *Econometrica*, **62**, 997–1040). The slightly weaker, but also easier to interpret conditions for the folk theorem given here were introduced by Kandori and Matsushima 1998 ("Private Observation, Communication and Collusion," *Econometrica*, **66**, 627–652) in an environment with private monitoring. Kandori 2003 ("Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring," *Econometrica*, **71**, 345–353) shows that, if players could in addition communicate, the folk theorem would hold under the same full-dimensionality assumptions as under perfect monitoring. The "technical" result due to Fan used in Section IV on linear inequalities is stated in "Systems of Linear Inequalities," in *Linear Inequalities and Related Systems*, H. W. Kuhn and A. W. Tucker, Ed., Paper 5. Annals of Mathematics Studies, Vol. 38, Princeton: Princeton Univ. Press.

Finally, Fudenberg, Levine, and Takahashi 2007 ("Perfect Public Equilibrium when Players are Patient," *Games and Economic Behavior*, **61**, 27–49) show how the scoring algorithm must be adjusted to account for the possibility that int $\mathcal{H} = \varnothing$, and give a general characterization that applies to that case as well. The dual characterization can be found in Hörner, Takahashi and Vieille 2014 ("On the Limit Perfect Public Equilibrium Payoff Set in Repeated and Stochastic Games," *Games and Economic Behavior*, **85**, 70–83). The famous example showing how PPE can be restrictive, mentioned in Section V, is due to Radner, Myerson and Maskin, 1986 ("An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," *Review of Economic Studies*, **53**, 59–69).

The impact of imperfect monitoring is studied in Kandori 1992 ("The use of information in repeated games with imperfect monitoring," *Review of Economic Studies*, **59**, 581–594). The rates of convergence are derived in Hörner and Takahashi (2015, "How fast do equilibrium payoff sets converge in repeated games?" working paper, Yale University).