# A Folk Theorem Sampler



## Recapitulation

Consider the prisoners' dilemma game of Figure 1. The unique Nash equilibrium (in fact, the dominance-solvable outcome) requires both players to Fink, which is Pareto dominated by both players cooperating (i.e. playing Mum). We often observe cooperation in the real world; what must we add to our model in order that cooperation become rational? Perhaps cooperation would be rational when players acknowledge that they are in a repeated relationship: that there is actually a sequence of stage games, where a player's behavior in a stage can be conditioned upon the treatment she has received from other players in the past.

© 1996 by Jim Ratliff, <jim@virtualperfection.com>, <http://virtualperfection.com/gametheory>.

| | $M$ | $F$ |
|---|---|---|
| $M$ | 1, 1 | −1, 2 |
| $F$ | 2, −1 | 0, 0 |

Figure 1: A Prisoners' Dilemma

Disconcertingly, when the prisoners' dilemma is repeated a commonly known finite number of times, a backwards induction argument implies that in the unique subgame-perfect equilibrium of this repeated game the players will still Fink in every period.[1,2] The key to sustaining cooperative behavior in this game is to let the stage game be infinitely repeated.[3] In particular, when we assumed that players' preferences over stage-game payoff streams could be represented by the discounted sum of the stage-game payoffs, we saw that cooperation by both players in every period was a subgame-perfect equilibrium of the infinitely repeated prisoners' dilemma, as long as the players were "sufficiently patient"—their discount factors were sufficiently close to one.[4]

Cooperation was achieved in the infinitely repeated prisoners' dilemma by having both players adopt "grim trigger" strategies: Each starts out playing the desired Mum and continues to choose this cooperative action as long as the other player has always played Mum as well. If her opponent ever plays Fink, though, she switches to the open-loop strategy of Finking in every period.

To see that these trigger strategies constitute a Nash equilibrium of the repeated game we analyze the tradeoff that a player faces when deciding whether to deviate from her appointed actions along the alleged equilibrium path: Playing Fink rather than Mum benefits the deviant in that period because she receives two instead of one.[5] However, because upon a defection her opponent would switch to the open-loop punishment phase of Fink every period, the deviant can do no better than accept zero in every remaining period. So the prospective deviant asks herself which is more desirable: receiving two today and zero forever afterward? or conforming and receiving one today and forever? The answer depends on her discount factor $\delta$. If she is very impatient, valuing the future very little, i.e. $\delta$ is close to zero, the single-period gain in the defecting period will outweigh the loss in later periods. If, on the other hand, she is more patient, receiving one rather than zero in every subsequent period outweighs the temptation of today grabbing two rather than one. We saw that in the infinitely repeated version of the game in Figure 1 cooperation in every period was a Nash equilibrium as long as each player's discount factor exceeded one-half.

---

1 This is Theorem 4 of the "Repeated Games." The result follows because there is a unique Nash-equilibrium payoff vector in the stage game. As we saw in the two-period example in "Repeated Games," it does not hold when the stage game has more than one Nash-equilibrium payoff vector. See also Benoît and Krishna [1985].

2 In fact, (Fink, Fink) in every period is also the unique Nash equilibrium of the finitely repeated prisoners' dilemma. This requires more reasoning and relies upon the fact that the stage-game Nash equilibrium yields the minmax payoff vector. See Sorin [1986: 156, Proposition 13] or Nalebuff [1988: 150, 153–154, Puzzle 3].

3 Or the game could be finitely repeated with probability one, but players are uncertain about how long the game will continue and in every period they assess a positive probability to the game continuing at least one more period. See the "Infinitely Repeated Games with Discounting" handout.

4 See the "Infinitely Repeated Games with Discounting" handout.

5 We analyze a deviation assuming that all other players choose the actions prescribed by the conjectured equilibrium strategies. In this case, then, a player contemplating deviation in some period assumes that her opponent will play Mum that period.

These trigger strategies constitute a subgame-perfect equilibrium of the repeated game, for the same range of discount factors, because their restrictions to any subgame form a Nash equilibrium of the subgame. In particular, in any subgame in which either player had ever deviated, the equilibrium strategies dictate that the players choose (Fink, Fink) every period. Because this is just an open-loop repetition of a stage-game Nash equilibrium, the restrictions of the repeated-game strategies to such a subgame constitute a Nash equilibrium of the subgame.[6,7]

# Introduction

So we have successfully shown how cooperation can arise in a repeated relationship where it could not in a “single-shot” context. This raises additional questions. Can other outcomes besides “everyone always cooperate” or “everyone always fink” (which remains equilibrium behavior for all discount factors) be sustained in equilibrium in the infinitely repeated prisoners’ dilemma? What about other games? What outcomes can be sustained in equilibrium when they are infinitely repeated? The original motivation for developing a theory of repeated games was to show that cooperative behavior was an equilibrium. The theoreticians were all too clever, for, as we will see, they showed that in many cases a huge multiplicity of even very “noncooperative” stage-game payoffs could be sustained on average as an equilibrium of the repeated game.

These findings are made precise in numerous *folk theorems*.[8,9] Each folk theorem considers a class of games and identifies a set of payoff vectors each of which can be supported by some equilibrium strategy profile. There are many folk theorems because there are many classes of games and different choices of equilibrium concept. For example, games may be repeated infinitely or only finitely many times. There are many different specifications of the repeated game payoffs. For example, there is the Cesaro limit of the means, the Abel limit (Aumann [1985: 210]), the overtaking criterion (Rubinstein [1979]) as well as the average discounted payoff, which we have adopted. They may be games of complete information or they might be characterized by one of many different specifications of incomplete information. Some folk theorems identify sets of payoff vectors which can be supported by Nash equilibria; of course, of more interest are those folk theorems which identify payoffs supported by subgame-perfect equilibria.

The strongest folk theorems are of the following loosely stated form: “Any *strictly individually rational* and *feasible* payoff vector of the stage game can be supported as a subgame-perfect equilibrium average payoff of the repeated game.” These statements often come with qualifications such as “for discount factors sufficiently close to 1” or, for finitely repeated games, “if repeated sufficiently many times.”

---

6 See Theorem 1 of the “Repeated Games” handout.

7 Their restrictions to these deviation subgames are Nash equilibria of these subgames for all discount factors. The qualification on the discount factor is inherited by the requirement that the strategy profile be a Nash equilibrium of the entire game.

8 They are so named because results of their type were widely believed by game theorists prior to published formal statements or proofs. Myerson [1991: 332] suggests that these be called *general feasibility* theorems. He correctly points out that “naming a theorem for the fact that it was once unpublished conveys no useful information to the uninitiated reader.” However, I wouldn’t hold my breath waiting for Myerson’s suggestion to become universally accepted. Game theorists are a terminologically stubborn lot.

9 There is a large folk-theorem literature. For an introduction to it see Fudenberg and Tirole [1991: 150–160].

First we will precisely define the terms *feasible* and *individually rational*. A payoff vector's individual rationality relies on the concept of a player's *minmax value*—a number useful for characterizing the worst punishment to which a deviating myopic player can be subjected in a single player—so we will define it as well. Then we will state and prove two folk theorems, one Nash and one perfect, which have the virtue of being relatively easy to prove because their proofs rely only on simple "grim trigger" strategies. Then we will prove a perfect folk theorem stronger than the first two using more complicated strategies.

# Infinitely repeated games with discounting

We'll call the stage game $G$ and interpret it to be a simultaneous-move matrix game which remains exactly the same through time. As usual we let the player set be $I=\{1,\ldots,n\}$. Each player has a pure action space $A_i$. The space of action profiles is $A=\mathsf{X}_{i\in I}A_i$. Each player has a von Neumann-Morgenstern utility function defined over the outcomes of $G$, $g_i$: $A\to\mathbb{R}$. We let $g$ be the $n$-tuple of players' stage-game payoff functions, i.e. $g=\mathsf{X}_{i\in I}g_i$ so that $g(a)=(g_1(a),\ldots,g_n(a))$.[10]

The stage game repeats each period, starting at $t=0$. At the conclusion of each period's play, the action profile which occurred is revealed to all the players. Combined with perfect recall, this allows a player to condition her current action on all earlier actions of her opponents. A repeated-game strategy $s_i\in S_i$ for player $i$ is a sequence $s_i=(s_i^0,s_i^1,\ldots)$, where each $s_i^t$ is a history-dependent stage-game strategy, $s_i^t$: $A^t\to A_i$. The history at time $t$, viz. $h^t\in A^t$, is the sequence of action profiles $h^t=(a^0,a^1,\ldots,a^{t-1})$.

We can think of the players as receiving their stage-game payoffs period-by-period, where the players discount future payoffs according to a common discount factor $\delta\in(0,1)$. We adopt the *average discounted payoff* representation of players' preferences over streams of stage-game payoffs. A player's average discounted payoff for an infinite stream of stage-game payoffs $v_i^\infty\equiv v_i^0,v_i^1,\ldots$ is

$$\Lambda(v_i^\infty,\delta)=(1-\delta)\sum_{t=0}^{\infty}\delta^t v_i^t. \tag{1}$$

(The $\delta$ may be suppressed as an argument if it is not a parameter of interesting variation.) This formulation normalizes the repeated-game payoffs to be "on the same scale" as the stage-game payoffs. We refer to the infinitely repeated game as $G^\infty$. Because the discount factor $\delta$ enters the players' utility functions, it is part of the definition of the game. When we need to emphasize the role of the discount factor we more explicitly refer to the game as $G^\infty(\delta)$.

Any repeated-game strategy profile $s$ generates a path $h^\infty=(a^0,a^1,\ldots)$, where each $a^t=s^t(h^t)$ and the histories are generated recursively by concatenation as $h^t=(h^{t-1};s^{t-1}(h^{t-1}))$. Player $i$'s repeated-game payoff to player $i$ along this path is, from (1),

10 See the "Infinitely Repeated Games with Discounting" handout. I want to reserve "$s$", "$S$", and "$u$" to refer to typical strategies, strategy spaces, and utility functions, respectively, in the repeated game. So I'm using "$a$", "$A$", and "$g$" for the corresponding stage-game entities.

$$u_i(s) = (1-\delta)\sum_{t=0}^{\infty} \delta^t g_i(a^t). \tag{2}$$

We denote the $n$-tuple of repeated-game payoff functions by $u = (u_1, \ldots, u_n)$. If we need to emphasize the role of the discount factor we will write $u_i(s;\, \delta)$.

A useful formula for computing the infinite sums we will encounter is

$$\sum_{t=T_1}^{T_2} \delta^t = \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta}, \tag{3}$$

which, in particular, is valid for $T_2 = \infty$.[11] If the payoffs $v_i^t$ a player receives are some constant payoff $v_i'$ for the first $t$ periods, viz. $0, 1, 2, \ldots, t-1$, and thereafter she receives a different constant payoff $v_i''$ in each period $t, t+1, t+2, \ldots$, the average discounted value of this payoff stream is[12]

$$\left(1-\delta^t\right) v_i' + \delta^t v_i'', \tag{4}$$

i.e. a convex combination of the two stage-game payoffs. If she receives $v_i'$ for the $t$ periods $0, 1, \ldots, t-1$ as before, then $v_i''$ only in period $t$, and $v_i'''$ every period thereafter, the average discounted value of this three-valued payoff stream is

$$(1-\delta^t) v_i' + \delta^t [(1-\delta) v_i'' + \delta v_i''']. \tag{5}$$

# The no one-stage improvement principle

In order to check whether a repeated-game strategy profile $s \in S$ is a subgame-perfect equilibrium of a repeated game, we in principle need to examine the restriction $s|_{h^t}$ of that strategy profile $s$ to every possible subgame defined by an arbitrary history $h^t$. In order to check whether $s|_{h^t}$ is a Nash equilibrium of the subgame, we need to check whether any player $i \in I$ has a profitable deviation from her part of $s|_{h^t}$. A deviation can be very complicated; it can involve a counter-to-specification action at any or all of an infinite number of histories. The set of possible such deviations can be huge.

Fortunately there is a result from dynamic programming which greatly simplifies our deviation-checking task. We need only check every possible deviation of an extremely simple class. If none of those deviations are profitable for a deviating player, then the strategy profile is subgame perfect. The simple deviations we consider are called "one-stage" deviations. In a one-stage deviation, the deviating player disobeys her component of the specified strategy profile at only a single history $h^t$. At all other information sets (i.e. histories), she obeys the prescription.

The intuition behind the result is loosely the following: The change in the repeated-game payoff due

11 See the "Appendix: Discounting Payoffs," in the "Infinitely Repeated Games with Discounting" handout for a derivation of this formula.

12 See the "Infinitely Repeated Games with Discounting" handout for the details.

to a complicated deviation from the prescribed strategy is a sum of the changes due to one-stage deviations. If a complicated deviation is to result in an increased repeated-game payoff, then some one-stage deviation must also result in an increased repeated-game payoff. If no one-stage deviation does, then no complicated deviation will be profitable either.

We say that the repeated-game strategy for player $i$, $\hat{s}_i \in S_i$, is a *one-stage deviant* of $s_i \in S_i$ if there exist a time $\hat{t}$ and a history $\hat{h}^{\hat{t}} \in A^{\hat{t}}$ such that 1 $\forall t \neq \hat{t}$, $\hat{s}_i^t = s_i^t$, 2 $\forall\ h^{\hat{t}} \in A^t \backslash \{\hat{h}^{\hat{t}}\}$, $\hat{s}_i^{\hat{t}}(h^{\hat{t}}) = s_i^{\hat{t}}(h^{\hat{t}})$, and 3 $\hat{s}_i^{\hat{t}}(\hat{h}^{\hat{t}}) \neq s_i^{\hat{t}}(\hat{h}^{\hat{t}})$. Let $\hat{S}_i(s_i)$ be the space of all one-stage deviants of $s_i$.

We say that $s$ satisfies the "no one-stage improvement (NOSI)" property if for all $i \in I$, for all one-stage deviants $\hat{s}_i \in \hat{S}_i(s_i)$ of $s_i$, for all $t$, and for all $h^t \in A^t$, $\hat{s}_i$ is no better than $s_i$ against $s_{-i}$, conditional on reaching the history $h^t$.

**Theorem** Let $s$ be a strategy profile for a finitely repeated game or an infinitely repeated game with discounting. Then $s$ is a subgame-perfect equilibrium if and only if $s$ satisfies the no one-stage improvement property.

**Proof** See Appendix.[13]

## The sufficient-patience limit

In proofs of the various folk theorems we'll be considering we will often be taking limits of average discounted payoffs as the discount rate tends toward unity. If there is a point in time beyond which a player receives some constant payoff for every period thereafter, such a limit is very simple: it is simply that constant payoff. For example, we calculated in (4) the average discounted payoff to a player if she received $v_i'$ for a finite number of periods and $v_i''$ for all remaining periods. It is clear in (4) that, in the limit as $\delta \to 1$, the first term vanishes leaving the limit to be the infinitely repeated payoff $v_i''$. A similar analysis shows that the limit of (5) is $v_i'''$.

We say that a payoff sequence $v_i^\infty \equiv v_i^0, v_i^1, \ldots$ has a *terminal subsequence* $\bar{v}_i$ if there exists a time period $\tau$ such that, for all $t > \tau$, $v_i^t = \bar{v}_i$.

**Lemma 1** Let $v_i^\infty \equiv v_i^0, v_i^1, \ldots$ be an infinite sequence of payoffs with terminal subsequence $\bar{v}_i$. Then

$$\lim_{\delta \to 1} \Lambda(v_i^\infty, \delta) = \bar{v}_i. \tag{6}$$

**Proof** The proof is left as an exercise for you. Hint: use (3). ☺

Sometimes we will consider two infinite sequences of payoffs, e.g. one for conforming and one for deviating, and we will wish to compare them in the infinite-patience limit. The following result can be

13 Warning! The proof I provide needs a lot of work to motivate it and make it more accessible.

very useful for performing this comparison.

**Theorem 1** Let $\hat{v}_i^\infty$ and $\check{v}_i^\infty$ be two infinite payoff sequences with terminal subsequences $\hat{v}_i$ and $\check{v}_i$, respectively. If $\hat{v}_i > \check{v}_i$, then, for a sufficiently patient player, the payoff stream $\hat{v}_i^\infty$ is strictly preferred to the payoff stream $\check{v}_i^\infty$. In other words, there exists a $\underline{\delta} \in (0, 1)$ such that, for all $\delta \in (\underline{\delta}, 1)$,

$$\Lambda(\hat{v}_i^\infty, \delta) > \Lambda(\check{v}_i^\infty, \delta).$$[14]

---

**Proof** (To simplify the notation, let's suppress the "$i$" subscripts.) Let $\varepsilon = \frac{1}{2}(\hat{v} - \check{v})$. From Lemma 1, we know that we can find $\delta', \delta'' \in (0, 1)$ such that, for all $\delta \in (\delta', 1)$,

$$|\ \hat{v} - \Lambda(\hat{v}^\infty, \delta)\ | < \varepsilon, \tag{♥.1}$$

and, for $\delta \in (\delta'', 1)$,

$$|\ \check{v} - \Lambda(\check{v}^\infty, \delta)\ | < \varepsilon. \tag{♥.2}$$

Let $\underline{\delta} = \max\,\{\delta', \delta''\}$. Then, for all $\delta \in (\underline{\delta}, 1)$,

$$\Lambda(\hat{v}^\infty, \delta) > \hat{v} - \varepsilon, \tag{♥.3}$$

$$\Lambda(\check{v}^\infty, \delta) < \check{v} + \varepsilon, \tag{♥.4}$$

from which we obtain

$$\Lambda(\check{v}^\infty, \delta) < \tfrac{1}{2}(\hat{v} + \check{v}) < \Lambda(\hat{v}^\infty, \delta), \tag{♥.5}$$

as desired. (See Figure 18.) ☺

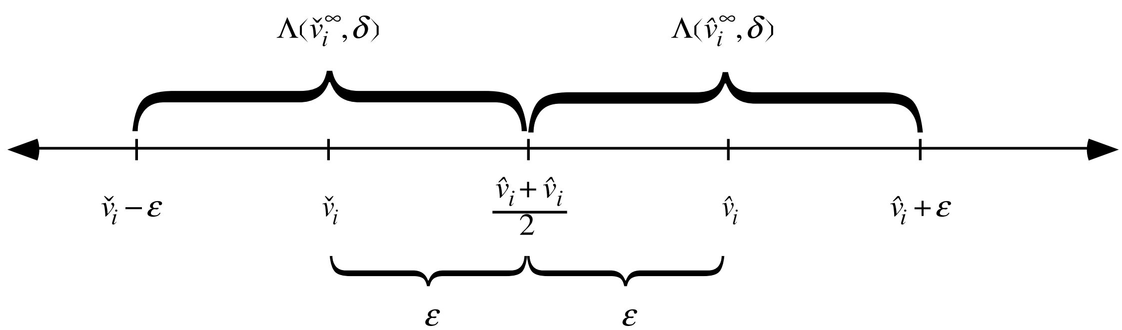


Figure 18: Taking $\delta$ sufficiently close to one ensures that the payoff with the higher terminal subsequence is more preferred.

---

14 Note that the weak-inequality form of this theorem is *not* true. In other words, if two payoff sequences have identical terminal subsequences, they need not have identical average discounted values, no matter how patient the player is. (However, the limit, as $\delta \to 1$, of the difference between the two average discounted values will go to zero.)

# Feasible stage-game payoffs

Say that we wish to support some payoff vector $v \in \mathbb{R}^n$ as an equilibrium of an infinitely repeated game with discounting. We must find an equilibrium strategy profile of the repeated game which generates an equilibrium-path history such that each player $i \in I$ receives a repeated-game payoff, from (2), of $v_i$.

We have already seen that the average discounted value of a constant stage-game payoff stream is independent of the discount factor $\delta$. An analytical inconvenience arises if the equilibrium path generates payoffs to player $i$ which change from period to period, i.e. when $g_i(a^t)$ varies with $t$. In this case, unlike the constant-stream case, the value of the average discounted value in (2) depends on the discount factor $\delta$. For example, say that the equilibrium path generates stage-game payoffs $v'$ in even numbered periods and $v''$ in odd numbered periods. The average discounted value of this stream is

$$(1-\delta)(v' + \delta v'' + \delta^2 v' + \delta^3 v'' + \cdots) = \frac{v' + \delta v''}{1+\delta}, \tag{7}$$

which strictly increases (respectively, decreases) with $\delta$ when $v'' > v'$ (respectively, $v' > v''$).

What's so problematic about the dependence on $\delta$ of the average discounted value? In arguments we will employ later, we fix an average-discounted payoff vector $v \in \mathbb{R}^n$ and take limits as $\delta \to 1$. If the payoff vector for a given equilibrium path changed with $\delta$, we would have to adjust the equilibrium strategies at every point in the limit so that, for each value of $\delta$, $v$ was achieved by those strategies. That would lead to a bona fide Excedrin™ intractability headache!

We can avoid this problem by limiting our attention to repeated-game strategy profiles whose equilibrium paths prescribe the same stage-game action profile in every period; i.e. $a^t = \bar{a}$ for all $t$. Then $g_i(a^t)$ is constant in time, and we know that the average-discounted value of this path for player $i$ is simply $g_i(\bar{a})$, which is conveniently independent of the discount factor $\delta$.

But at what cost is this proposed restriction of attention to strategy profiles which prescribe exactly the same stage-game profile in every period along the equilibrium path? Is this without loss of generality? In other words might there not be payoff vectors which could be sustained by time-varying equilibrium paths but which cannot be sustained by time-independent paths? The answer is: yes, this could be the case. We will next see the origin of the problem, and then we will augment our model with a public randomizing device in order to restore generality to our results. Of course, this augmentation has its own cost—what if the economic situation being studied does not admit such a device? (See Fudenberg and Tirole [1991: 152] for references to research which attacks this issue.)

## Stage-game payoffs with uncorrelated randomization

Consider the game of Figure 2. The game's three distinct pure-strategy payoff vectors are plotted in "payoff space" in Figure 3. The union of the two shaded regions is the *convex hull* of these three vectors.[15] Note that the $(0,1) \to (1,0)$ edge can be achieved by player 1 choosing Down and player 2

15 Let $X$ be a finite set of $m$ vectors in a real topological linear space $Y$ (e.g., $Y = \mathbb{R}^n$), $X = \{x^1, \ldots, x^m\} \subset Y$, and let $\Delta^k$ be the $k$-dimensional unit simplex. The *convex hull* of $X$, viz. $\operatorname{co} X$, is the set of vectors in $Y$ each of which is a convex combination of the vectors in $X$. In other words

mixing between left and right. It can also be achieved by Column choosing left and Row mixing between Up and Down. Similarly, the $(0,0)\to(1,0)$ edge can be achieved in either of two ways: 1 by player 1 choosing Up and player 2 mixing between right and left and 2 player 2 choosing right and player 1 mixing between Up and Down. In these cases each pair of adjacent extreme points of the convex hull corresponds to a pair of strategy profiles such that one player chooses the same pure strategy and one player chooses different pure strategies.[16]

However, this is not the case for the $(0,0)\to(0,1)$ edge. The $(0,0)$ payoff vector is obtained only from the strategy profile $(U,r)$, and the $(0,1)$ vector is obtained only from $(D,l)$. In this case each player chooses different pure strategies in the two strategy profiles.

$$\operatorname{co} X=\left\{y\in Y\colon \exists\beta\in\Delta^{m-1},\ \text{such that } \textstyle\sum_{i=1}^{m}\beta_i x^i=y\right\}.$$

16 Let $X$ be a convex set. A point $x\in X$ is an *extreme point* of $X$ if $x$ cannot be expressed as the midpoint of two distinct points in $X$; i.e. if $\nexists y,z\in X$ such that $y\neq z$ and $x=(y+z)/2$.

| | $l$: $[q]$ | $r$: $[1-q]$ |
|---|---|---|
| $U$: $[p]$ | 1,0 | 0,0 |
| $D$: $[1-p]$ | 0,1 | 1,0 |

Figure 2: A game whose pure-strategy payoffs′ convex hull cannot be achieved with uncorrelated mixed strategies.

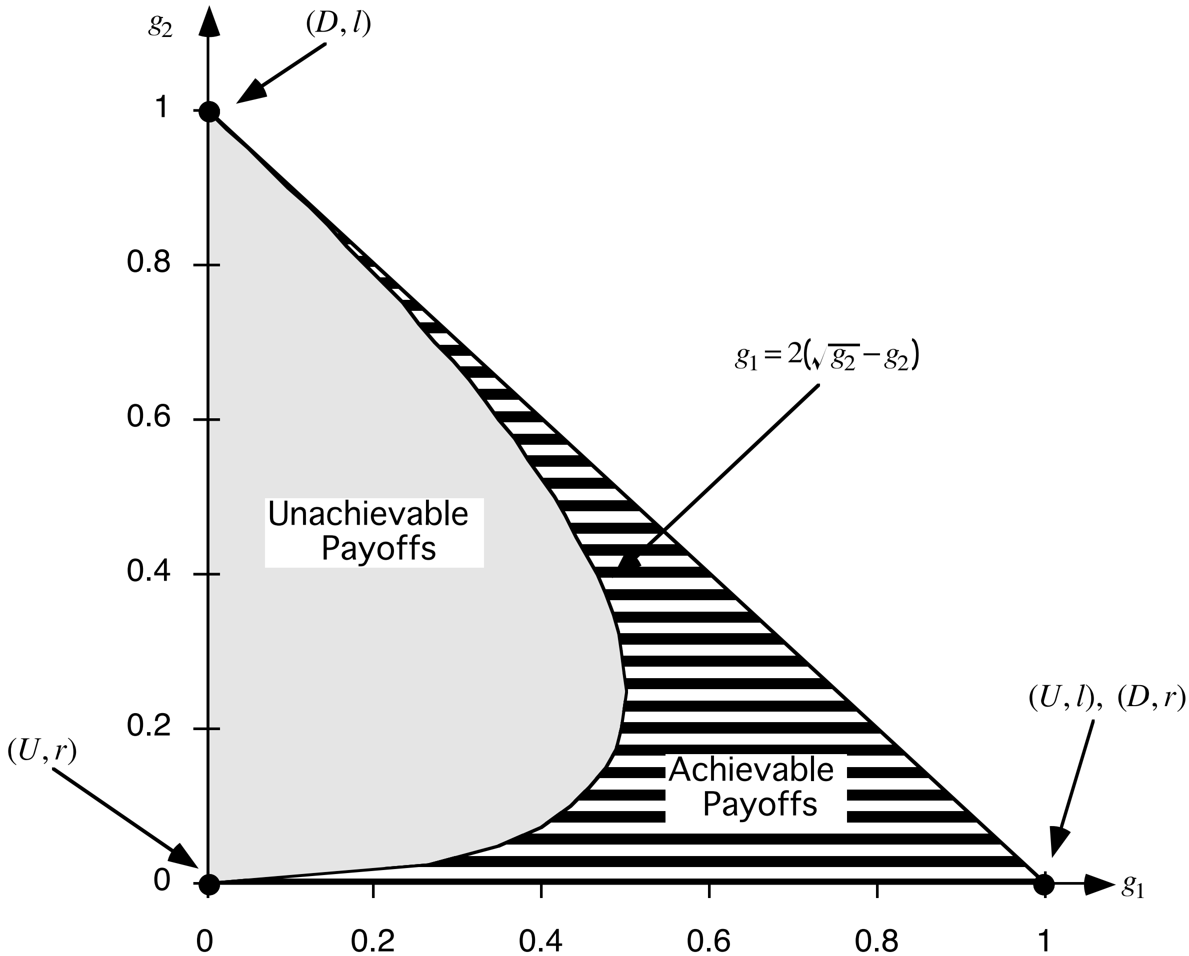


Figure 3: The payoff space for the game of Figure 1, indicating the unachievable and achievable subsets of the convex hull of the pure-strategy payoffs.

It is immediate to see why this edge cannot be achieved by uncorrelated randomizations. This edge is characterized by $g_1 = 0$. The only way such a payoff can be assured by the first player is if the only two strategy profiles receiving positive probability are $(D, l)$ and $(U, r)$. If a point $(0, v_2)$, $v_2 \in (0, 1)$, is to be achieved, player 1 or player 2 must strictly mix. (The only values of $g_2$ obtained by pure-strategy profiles are 0 and 1.) Such a strategy profile always results in positive weight applied to a strategy profile which yields player 1 a payoff of 1. Therefore an expected payoff for player 1 of zero cannot be obtained for $v_2 \in (0, 1)$.[17]

## Public randomization device

We saw that in the game of Figure 2 we could not achieve any payoff vector of the form $(0, v_2)$ where $v_2 \in (0, 1)$ when both players were limited to independently randomized mixed strategies. So we now introduce a publicly observable randomization device, which will allow the players to correlate their randomized actions. Specifically, assume that there exists a unit roulette wheel which, when spun, lands on a number between zero and one. Its pointer can be considered a random variable $\omega$ which is uniformly distributed on $[0, 1]$. It is perched atop A-Mountain and therefore visible to all players. Before each period, the wheel is spun and all players observe the realization of $\omega$ before they choose their actions for that period. Because $\omega$ becomes part of all players' information, they can condition their actions upon $\omega$.

For example, we could achieve the expected payoff vector $(0, v_2)$ if we could coax the players into playing the action profile $(D, l)$ with probability $v_2$ and the action profile $(U, r)$ with probability $(1 - v_2)$. The expected payoff vector from such correlated actions is, as desired,

$$g(v_2 \circ (D, l) \oplus (1 - v_2) \circ (U, r)) = v_2 g((D, l)) + (1 - v_2) g((U, r)) = v_2(0, 1) + (1 - v_2)(0, 0) = (0, v_2). \tag{8}$$

In order to implement this coordination we assign the players the following strategies:

$$a_R(\omega) = \begin{cases} D, & \omega \in [0, v_2), \\ U, & \omega \in [v_2, 1], \end{cases} \qquad a_C(\omega) = \begin{cases} l, & \omega \in [0, v_2), \\ r, & \omega \in [v_2, 1]. \end{cases} \tag{9}$$

Because $\omega$ is uniformly distributed on the unit interval, the probability that the pointer takes on a value in some interval $\langle c, d \rangle \subset [0, 1]$ is simply $d - c$.[18] Therefore the probability that $\omega \in [0, v_2)$ is $v_2$. When this event occurs, Row will choose $D$ and Column will choose $l$; therefore the action profile $(D, l)$ will occur with probability $v_2$. Similarly, $(U, r)$ will occur with probability $(1 - v_2)$. Therefore the pair of stage-game strategies in (9) results in the expected payoff vector $(0, v_2)$ as calculated in (8), even when $v_2 \in (0, 1)$.

More generally, consider the set of pure-action payoff vectors $g(A) \subset \mathbb{R}^n$.[19] Let $g(A) = \{\tilde{v}^1, \tilde{v}^2, \ldots, \tilde{v}^m\}$, where $m = \#g(A)$.[20] Each payoff vector $\tilde{v}^k$ can be achieved by some pure-action profile $\tilde{a}^k \in A$; i.e. $g(\tilde{a}^k) = \tilde{v}^k$. Let $V$ be the convex hull of these pure-action payoff vectors; i.e.

$$V = \operatorname{co} g(A) = \left\{ v \in \mathbb{R}^n \colon \exists \beta \in \Delta^{m-1},\ v = \sum_{k=1}^{m} \beta_k \tilde{v}^k \right\}. \tag{10}$$

To achieve any vector $v \in V$ as an expected payoff vector of the players' stage-game correlated-action

---

17 You can show that the left-most frontier of the achievable payoffs in Figure 3 is defined by $\{(g_1, g_2) \colon g_1 = 2(\sqrt{g_2} - g_2),$Ù$g_2 \in [0, 1]\}$.

18 Interpret the "⟨" as being either "(" or "[" and similarly for "⟩". In other words, because the probability distribution is continuous, it doesn't matter whether the interval includes none, one, or both of its endpoints.

19 Let $f \colon X \to Y$. The *image of X under f*, $f(X)$, is the set of values in the range of $f$ which are achieved for some value in the domain. I.e. $f(X) = \{y \in Y \colon \exists x \in X, f(x) = y\}$.

20 Note that $\#g(A) \leq \#A$.

profile, we find a corresponding vector $\beta$ of convex coefficients and assign each player $i$ the stage-game strategy

$$a_i(\omega)=\begin{cases} \tilde{a}_i^1, & \omega\in[0,\beta_1), \\ \tilde{a}_i^2, & \omega\in[\beta_1,\beta_1+\beta_2), \\ \tilde{a}_i^3, & \omega\in[\beta_1+\beta_2,\beta_1+\beta_2+\beta_3), \\ \vdots & \vdots \\ \tilde{a}_i^k, & \omega\in\left[\sum_{j=1}^{k-1}\beta_j,\sum_{j=1}^{k}\beta_j\right), \\ \vdots & \vdots \\ \tilde{a}_i^m, & \omega\in[\beta_1+\cdots+\beta_{m-1},1]. \end{cases} \tag{11}$$

For example, if the roulette wheel is spun and the realization $\omega$ lies in the interval $(\beta_1+\cdots+\beta_{k-1},\beta_1+\cdots+\beta_k)$, then each player $i\in I$ will choose $\tilde{a}_i{}^k$; i.e. the action profile $\tilde{a}^{\tilde{k}}$ will be chosen by the players. We note that the probability that any given realization of $\omega$ falls in the interval corresponding to the action profile $\tilde{a}^k$, and thus to the payoff vector $\tilde{v}^k$, is, as desired,

$$\sum_{j=1}^{k}\beta_j-\sum_{j=1}^{k-1}\beta_j=\beta_k. \tag{12}$$

Because every payoff in the convex hull of the pure-action payoffs can be achieved by some correlated action profile $a(\omega)$, we call $V$ the set of *feasible* outcomes.

## $\omega$-augmented histories

Players can condition their current action upon everything they know. Before we introduced the public randomization device, we assumed the only knowledge players gained in the course of play was their observations of other players' past pure actions. Now they also learn in each period the realization $\omega^t$ of the random variable $\omega$ in that period. When a player is choosing her period-$t$ action, then she knows all of the action profiles played in previous periods as well as all realizations of $\omega$ in previous periods *and* in the current period. (The current period's realization is revealed prior to taking that period's action.) Therefore we write the history in period $t$ as the $(2t+1)$-tuple

$$h^t=(a^0,a^1,\ldots,a^{t-1};\omega^0,\omega^1,\ldots,\omega^t). \tag{13}$$

Before we performed this augmentation, the space of period-$t$ histories was simply $A^t$, the $t$-fold Cartesian product of action-profile spaces. Now it is more complicated. Each $\omega^t$ is a number on the unit interval. Therefore the space of augmented histories is

$$H^t=A^t\times[0,1]^{t+1}. \tag{14}$$

A repeated-game strategy for player $i$ is still a sequence $s_i=(s_i^0,s_i^1,\ldots)$, but now each history-dependent stage-game strategy in the sequence is a function $s_i^t\colon H^t\to A_i$, where $H^t$ is as defined in (14).

# Minmax punishments

In order to implement some payoff vector $v \in V$ as an average discounted payoff of a repeated-game equilibrium, we will first find a correlated-action profile $a(\omega)$ whose expected payoff is $v$ (where the expectation is taken with respect to $\omega$); i.e.

$$\mathbb{E}_\omega g(a(\omega)) = v. \tag{15}$$

In equilibrium, then, in every period $t$ we want each player $i$ to choose the action $a_i(\omega^t)$, where $\omega^t$ is the realization of the random variable $\omega$ in period $t$.

However, we can't necessarily just prescribe that each player choose $a(\omega)$ every period regardless of what her opponents do. In other words, we can't necessarily prescribe open-loop strategies to the players. The reason is that it need not be the case that for all $\omega$ (or even for any $\omega$) the action profile $a(\omega)$ is a Nash equilibrium of the stage game. Therefore there could be some $\omega$ for which some player $i$ would receive a higher stage-game payoff by playing some $\hat{a}_i \in A_i$ rather than $a_i(\omega)$; i.e. $g_i(\hat{a}_i, a_{-i}(\omega)) > g_i(a(\omega))$. If her opponents were playing open-loop strategies, she could deviate to this $\hat{a}_i$ with impunity because it would not affect her opponents' future actions. Therefore the open-loop prescription of $a(\omega)$ every period would not be an equilibrium of the repeated game.

To encourage each player $i$ to play her part $a_i(\omega)$ of the equilibrium-path prescription, the other players must plan to punish her if she deviates. It will be useful to identify the most severe punishment with which a player can be inflicted. Consider the case where all players other than $i$ set out to punish player $i$ as severely as they can.

We will allow the punishers to choose mixed stage-game actions, so we denote by $\mathscr{A}_j \equiv \Delta(A_j)$ the space of player-$j$ mixed actions in the stage game.[21] A typical mixed action for player $j$ is $\alpha_j \in \mathscr{A}_j$. A typical deleted mixed-action profile by the punishers is $\alpha_{-i} \in \mathscr{A}_{-i} \equiv \times_{j \in I \setminus \{i\}} \mathscr{A}_j$.

## The minmax vector

In any equilibrium all players know the strategies of the other players. Therefore whatever punishing actions $\alpha_{-i}$ the punishers choose, player $i$ knows $\alpha_{-i}$ and will play a best response to it. Player $i$ will therefore receive a payoff $\bar{g}_i(\alpha_{-i})$ defined by

$$\bar{g}_i(\alpha_{-i}) \equiv \max_{a_i \in A_i} g_i(a_i, \alpha_{-i}). \tag{16}$$

[Make sure you understand why it is without loss of generality that we can restrict attention to player-$i$'s pure-strategy space in solving (16).] Her punishers know that player $i$ will choose $a_i$ which solves (16), so they choose their deleted mixed-action punishment profile $m_{-i}^i$ to minimize the payoff player $i$ receives given that she optimizes against the punishment, i.e. to solve

21 "Ç($X$)" denotes the set of all probability distributions over the finite set $X$.

$$m_{-i}^i \in \underset{\alpha_{-i}\in\mathscr{A}_{-i}}{\arg\min}\left(\max_{a_i\in A_i} g_i(a_i,\alpha_{-i})\right) = \underset{\alpha_{-i}\in\mathscr{A}_{-i}}{\arg\min}\ \bar{g}_i(\alpha_{-i}).^{22} \tag{17}$$

Player $i$'s best response to $m_{-i}^i$, then, is some $m_i^i \in A_i$ satisfying

$$m_i^i \in \underset{a_i\in A_i}{\arg\max}\ g_i(a_i, m_{-i}^i), \tag{18}$$

where we have substituted $m_{-i}^i$ for $\alpha_{-i}$ in (16). The mixed-action profile $m^i=(m_i^i, m_{-i}^i)$ is player $i$'s *minmax vector*. For any $j\in I$, $m_j^i$ is player $j$'s component of the minmax mixed-action stage-game action profile which punishes player $i$ as severely as possible (given that player $i$ is myopically resisting the punishment). Note that the minmax mixed-action profile is a strategic specification for both the punished and the punishers. Player $i$'s *minmax payoff* $\underline{v}_i$ is the lowest expected stage-game payoff to which she can be held by players trying to punish her as severely as possible; this value is

$$\underline{v}_i = g_i(m^i) = \bar{g}_i(m_{-i}^i) = \min_{\alpha_{-i}\in\mathscr{A}_{-i}}\left(\max_{a_i\in A_i} g_i(a_i,\alpha_{-i})\right) = \min_{\alpha_{-i}\in\mathscr{A}_{-i}} \bar{g}_i(\alpha_{-i}). \tag{19}$$

We denote the $n$-tuple of the players' minmax values, the *minmax vector*, by $\underline{v}=(\underline{v}_1,\ldots,\underline{v}_n)$. (The minmax vector need not be feasible; i.e. it could be the case that $\underline{v}\notin V$.)

**Example: Calculating the minmax vector in a $3\times 2$ game**

Consider the two-player, $3\times 2$ stage-game below.

| | $l$: $[r]$ | $r$: $[1-r]$ |
|---|---|---|
| $U$: $[p]$ | $-2$, 2 | 1, $-2$ |
| $M$: $[q]$ | 1, $-2$ | $-2$, 2 |
| $D$: $[1-p-q]$ | 0, 1 | 0, 1 |

First let's look for each player's pure-action minmax value—the lowest payoff to which she can be held if her opponents can only play pure actions. To find Row's pure-action minmax value $\underline{v}_R^p$, we consider each of Column's pure strategies, finding Row's best-response payoff for each. We have

$$\bar{g}_R(l)=1 \quad\text{and}\quad \bar{g}_R(r)=1.$$

Therefore Row's pure-action minmax payoff is

$$\underline{v}_R^p = \min\{\bar{g}_R(l), \bar{g}_R(r)\} = 1.$$

Similarly, Column's pure-action minmax payoff is

$$\underline{v}_C^p = \min\{\bar{g}_C(U), \bar{g}_C(M), \bar{g}_C(D)\} = \min\{2,1,1\} = 1.$$

Therefore the players' pure-action minmax payoff vector is $\underline{v}^p=(1,1)$. We will see that their mixed-action minmax payoff vector is strictly lower.

22 The occurrence of "min max" explains the origin of the term minmax.

Let the probability with which Column chooses $l$ be $t$, and let the probabilities with which Row chooses $U$ and $M$ be $p$ and $q$, respectively.

To find Row's minmax value $\underline{v}_R$ we compute Row's expected utility for each of her pure actions as a function of Column's mixed action $t$. The upper envelope of these three functions of $t$ is itself a function of $t$ and represents Row's expected utility when she maximizes against Column's action. The minimum value of this upper envelope will be Row's minmax value. Any minimizer (i.e. a value of $t$) of this upper envelope will be a punish-Row minmax deleted action profile $m_C^R$. We have

$$u_R(U;t) = -2t + (1-t) = 1 - 3t,$$

$$u_R(M;t) = t - 2(1-t) = -2 + 3t,$$

$$u_R(D;t) = 0.$$

See Figure Ex. 1.

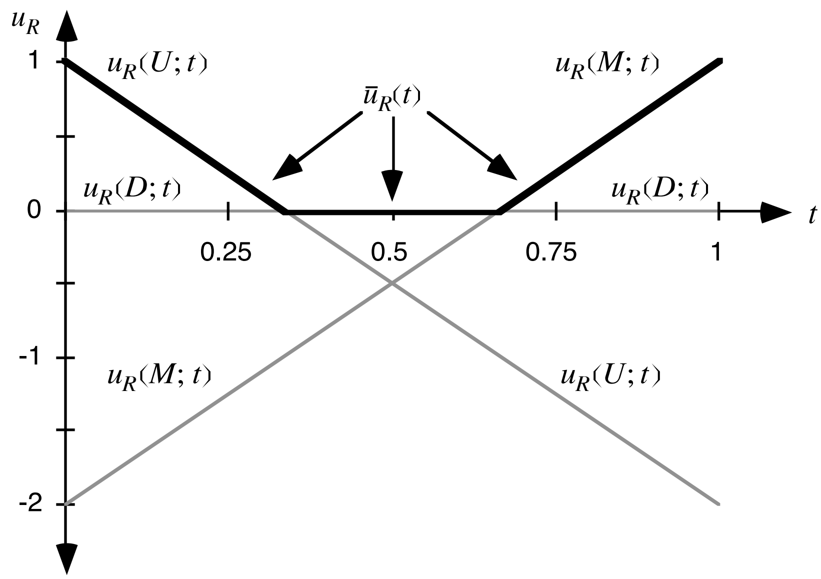


Figure Ex. 1: Row's expected payoffs to her pure actions as a function of Column's mixed action $t$.

The upper envelope of these functions is

$$\bar{u}_R(t) = \max\{u_R(U;t), u_R(M;t), u_R(D;t)\} = \begin{cases} 1-3t, & t\in[0, 1/3), \\ 0, & t\in[1/3, 2/3], \\ -2+3t, & t\in(2/3, 1]. \end{cases}$$

The minimum value of this upper envelope is zero, which is achieved for any $t\in[\frac{1}{3}, \frac{2}{3}]$. Therefore Row's minmax payoff is $\underline{v}_R = 0$. Therefore one minmax punish-Row deleted action profile would be $m_C^R = \frac{1}{2}\circ l \oplus \frac{1}{2}\circ r$. Row's best response to this mixed action of Columns is $m_C^C = D$.

To find Column's minmax value $\underline{v}_C$ we compute Column's expected utility for each of her two pure actions as a function of Row's mixing probabilities $p$ and $q$. We have

$$u_C(l;p,q)=2p-2q+(1-p-q)=1+p-3q,$$

$$u_C(r;p,q)=-2p+2q+(1-p-q)=1-3p+q.$$

To determine the upper envelope of this function we note that $u_C(l;p,q)\geq u_R(r;p,q)$ if and only if $p\geq q$. Let $\Delta\equiv\{(p,q)\in\mathbb{R}_+{}^2 : p+q\leq 1\}$, $\Delta^l=\Delta\cap\{(x,y)\in\mathbb{R}^2 : x\geq y\}$, and $\Delta^r=\Delta\cap\{(x,y)\in\mathbb{R}^2 : x<y\}$. (See Figure Ex. 2.) The two triangles $\Delta^l$ and $\Delta^r$ form a partition of $\Delta$ such that 1 when restricted to $\Delta^l$, $u_C(l;p,q)\geq u_C(r;p,q)$ and 2 when restricted to $\Delta^r$, $u_C(r;p,q)>u_C(l;p,q)$. Then

$$\bar{u}_C(p,q)=\max\{u_C(l;p,q),u_C(r;p,q)\}=\begin{cases}1+p-3q, & (p,q)\in\Delta^l,\\ 1-3p+q, & (p,q)\in\Delta^r.\end{cases}$$

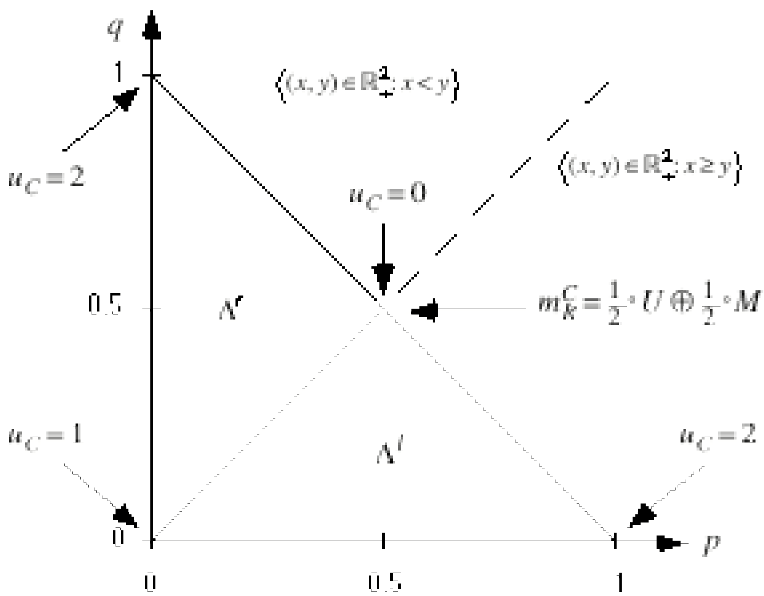


Figure Ex. 2: Row's mixed-actions as points in a two-dimensional simplex.

To find the minimum of $\bar{u}_C$ over $\Delta$ we can simply find the minimum of $\bar{u}_C$ over each cell of the partition and then take the minimum of these two minimum values. I.e.

$$\min_{(p,q)\in\Delta}\bar{u}_C(p,q)=\min\left\{\min_{(p,q)\in\Delta^l}\bar{u}_C(p,q),\ \min_{(p,q)\in\Delta^r}\bar{u}_C(p,q)\right\}.$$

Each of the simple minimization problems, viz. $\min\{\bar{u}_C(p,q) : (p,q)\in\Delta^a\}$ for $a\in\{l,r\}$, is simply a minimization of a linear function over a convex polyhedron.[23] The minimum value must occur at a vertex of the convex polyhedron.[24] The vertices of $\Delta^l$ are $(0,0)$, $(1,0)$, and $(\frac{1}{2},\frac{1}{2})$. The values of $\bar{u}_C$ at these vertices are 1, 2, and 0, respectively. Therefore the minimum of $\bar{u}_C$ over $\Delta^l$ is 0, which is achieved

23 A set is a *convex polyhedron* if it is the convex hull of a finite set.

24 See Intriligator [1971: 75].

at $(\frac{1}{2},\frac{1}{2})$. The vertices of $\Delta^r$ are $(0,0)$, $(0,1)$, and $(\frac{1}{2},\frac{1}{2})$. The values of $\Delta^r$ at these vertices are 1, 2, and 0, respectively. Therefore the minimum of $\bar{u}_C$ over $\Delta^r$ is 0 which is achieved at $(\frac{1}{2},\frac{1}{2})$. Therefore the minimum of $\bar{u}_C$ over $\Delta$ is 0, and is achieved at $(\frac{1}{2},\frac{1}{2})$.

Therefore Column's minmax payoff is $\underline{v}_C=0$. The minmax punish-Column deleted action profile then is $m_R{}^C=\frac{1}{2}\circ U\oplus\frac{1}{2}\circ M$. Column is indifferent between $l$ and $r$ against this $m_R{}^C$, so we can arbitrarily choose any Column mixed action for $m_C{}^C$.

Therefore the minmax payoff vector is $\underline{v}=(0,0)$, which is strictly less than the pure-action minmax payoff vector; i.e. $\underline{v}\ll\underline{v}^p$.

## Individual rationality

A player's minmax value $\underline{v}_i$ establishes a lower bound for the payoff she can be forced to receive in any stage-game Nash equilibrium as well as for the average discounted payoff she can receive in any repeated-game equilibrium. More formally….

**Lemma 2** Let $\bar{\alpha}\in\mathscr{A}$ be a stage-game mixed-action profile such that $\bar{\alpha}_i$ is a best response by player $i$ to $\bar{\alpha}_{-i}$. Then, $g_i(\bar{\alpha})\geq\underline{v}_i$.

**Proof** Because $\bar{\alpha}_i$ is a best response to $\bar{\alpha}_{-i}$,

$$g_i(\bar{\alpha})=\bar{g}_i(\bar{\alpha}_{-i})\geq\min_{\alpha_{-i}\in\mathscr{A}_{-i}}\bar{g}_i(\alpha_{-i})=\underline{v}_i,$$

where we have used (16), the definition of the minimum, and (19).

**Theorem 2** Let $\bar{\alpha}\in\mathscr{A}$ be a stage-game Nash-equilibrium mixed-action profile. Then, for all $i\in I$, $g_i(\bar{\alpha})\geq\underline{v}_i$, or equivalently, $g(\bar{a})\geq\underline{v}$.

**Proof** For every $i\in I$, $\bar{\alpha}_i$ must be a best response to $\bar{\alpha}_{-i}$; therefore the conclusion follows immediately from Lemma 2. ☺

**Theorem 3** Let $\bar{\sigma}$ be a Nash-equilibrium mixed-strategy profile of an infinitely repeated game with discounting. Then, for every $i\in I$, $u_i(\bar{\sigma})\geq\underline{v}_i$.

**Proof** Because $\bar{\sigma}$ is a Nash equilibrium, for all $i\in I$, $\bar{\sigma}_i$ is a best response to $\bar{\sigma}_{-i}$. Therefore for all player-$i$ repeated-game strategies $s_i\in S_i$,

$$u_i(\bar{\sigma})\geq u_i(s_i,\bar{\sigma}_{-i}). \tag{⑨.1}$$

Let $\mathsf{BR}_i\colon\mathscr{A}_{-i}\twoheadrightarrow A_i$, be player $i$'s stage-game pure-action best-response correspondence. Consider the player-$i$ repeated-game strategy $s_i=(s_i^0,s_i^1,\ldots)$, where in each period player-$i$ chooses a myopic best response to the actions of the other players, $\bar{\sigma}_{-i}^t(h^t)\in\mathscr{A}_{-i}$. I.e. for all $t\in\{0,1,\ldots\}$ and all $h^t\in A^t$,

$$s_i^t(h^t)\in \mathsf{BR}_i(\bar{\sigma}_{-i}^t(h^t)). \tag{⑨.2}$$

From Lemma 2 we know that, for all $t$ and $h^t$,

$$g_i(s_i^t(h^t),\bar{\sigma}_{-i}^t(h^t))=\bar{g}_i(\bar{\sigma}_{-i}^t(h^t)\geq \underline{v}_i. \tag{⑨.3}$$

Player $i$'s repeated-game payoff is

$$u_i(s_i,\bar{\sigma}_{-i})=(1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(s_i^t(h^t),\bar{\sigma}_{-i}^t(h^t))\geq(1-\delta)\sum_{t=0}^{\infty}\delta^t\underline{v}_i=\underline{v}_i, \tag{⑨.4}$$

where the $h^t$ are computed recursively in the usual fashion. The conclusion follows from (⑨.1). ☺

---

Theorem 3 tells us that it would be hopeless to try to sustain a payoff vector $v$ as an equilibrium of the repeated game if this would imply that some player were receiving less than her minmax value; i.e. that $\exists i\in I$ such that $v_i<\underline{v}_i$. Further, such a player would not even be optimizing against the strategies of her opponents. So we say that a payoff vector $v$ is *individually rational* if $v\geq\underline{v}$; i.e. if for all $i\in I$, $v_i\geq\underline{v}_i$. We say that $v$ is strictly individually rational if $v\gg\underline{v}$; i.e. for all $i\in I$, $v_i>\underline{v}_i$.[25] We denote by $V^*$ the intersection of the set $V$ of feasible payoffs and the set of strictly individually rational payoffs, i.e.

$$V^*=\{v\in V:\ v\gg\underline{v}\}. \tag{20}$$

# Grim trigger strategy folk theorems

We now present two folk theorems, which have the virtue of being particularly easy to prove. The strategies they employ to sustain a given payoff vector as an equilibrium have the following form: Play begins in the "normal phase," in which all players repeat every period their part of a stage-game action profile which achieves $v$. If a single player deviates from the normal phase, play switches to an open-loop, punish-the-defector phase. If players are sufficiently patient, the threat of the punishment phase will deter them from yielding to the temptation of normal-phase defection.

Adopting trigger strategies provides an important analytical simplification. The punishment phase kicks in when a deviation occurs; however, the details of the punishment phase depend only on the identity of the defector; they are independent of which particular deviation the deviator commits and independent of when she commits it. Therefore we can just consider the myopically best single-period deviation, because the punishment will be the same no matter what.

The first theorem chooses an arbitrary feasible and strictly individually rational payoff vector $v\in V^*$ and uses minmax threats to enforce a Nash-equilibrium path which yields the players an expected payoff

25 My terminology is somewhat at variance with the published literature, I'm afraid. What I refer to as the strictly individually rational set is usually called simply the individually rational set. However, there is nothing necessarily irrational about a player receiving exactly her minmax payoff rather than something strictly greater than that. My terminology also highlights a theoretical lacuna: The folk theorems we will study at most show that the strictly individually rational payoffs are equilibria. However, the larger set of (weakly) individually rational payoffs cannot be ruled out as equilibria. Therefore the folk theorems do not necessarily fully characterize the set of equilibrium payoffs.

of $v$. However, as we have already observed, the minmax action profile need not be a Nash equilibrium of the stage game. Therefore if a player does defect from the normal phase, the equilibrium strategies can prescribe an open-loop repetition of a stage-game action profile which is not a Nash equilibrium of the stage game. This behavior will not be a Nash equilibrium of that subgame.[26] Therefore this theorem will establish that the chosen payoff vector $v$ is a Nash-equilibrium payoff but will not show that it is a subgame-perfect equilibrium payoff. (We will later see a stronger folk theorem, which uses strategies more complicated than grim trigger strategies, which will show that these same feasible and strictly individually rational payoffs are subgame perfect.)

The second theorem focuses on a subset of the feasible and strictly individually rational payoffs $V^*$. We consider those payoffs which strictly Pareto dominate some stage-game Nash equilibrium payoff. (We actually look at a somewhat larger set, but I'll be more precise later.) Rather than punishing with the minmax strategy profile, if a player defects from the normal phase, play switches to the open-loop repetition of a stage-game Nash equilibrium which is worse for the deviator than her equilibrium payoff. The disadvantage of using stage-game Nash equilibria as the punishment profiles is that this method supports only a possibly proper subset of $V^*$ as equilibrium payoffs. The advantage is that now the punishment phase is a repeated-game Nash equilibrium in any subgame in which a player previously defected; therefore this subset of $V^*$ is shown to be composed not only of Nash-equilibrium payoffs but, more strongly, of subgame-perfect payoffs.

## Grim trigger strategies defined

Consider a payoff vector $v \in V^* \subset V = \operatorname{co} g(A)$. We know that there exists a correlated stage-game action profile $\bar{a}(\omega)$ which realizes the payoff vector $v$ in expectation over the random variable $\omega$. In every period $t$ of the "normal phase" of the game we want each player $i$ to choose her part of $\bar{a}(\omega^t)$. We say that player $j$ was the *solo deviator* in period $t$ if she alone did not play her part of $\bar{a}(\omega^t)$, i.e. if the actual stage-game action profile played in that period, $a^t \in A$, was such that $a^t_{-j} = \bar{a}_{-j}(\omega^t)$ but $a^t_j \neq \bar{a}_j(\omega^t)$. We say that player $j$ was the *earliest solo deviator* if in some period $t'$ player $j$ was the solo deviator and, for all prior periods $t'' \in \{0, 1, \ldots, t'-1\}$, there was no player $i \in I$ who was a solo deviator in period $t''$.

To be very formal… for any period $t$, define the earliest-solo-deviator function $\hat{\jmath}^t\colon H^t \to (I \cup \{0\})$, so that $\hat{\jmath}^t(h^t)$ is the earliest solo deviator, if any, according to the augmented history $h^t$. If no player ever deviated solo in periods $0, 1, \ldots, t-1$, set $\hat{\jmath}^t(h^t) = 0$. You can see that $\hat{\jmath}^t(h^0) = 0$, i.e. in period zero no player has previously deviated. Further, once a player has earned the distinction of being the earliest solo deviator, she always has that distinction. In other words, if $\hat{\jmath}^t(h^t) \in I$, then for every extension $h^{t+\tau} = (h^t; a^{t+1}, \ldots, a^{t+\tau-1})$ of $h^t$, $\hat{\jmath}^{t+\tau}(h^{t+\tau}) = \hat{\jmath}^t(h^t)$.

Consider a set of $n$ mixed-action profiles $b^i \in \mathscr{A}$, $i \in I$. The interpretation is that $b^i$ is a stage-game profile intended to punish player $i$ if she deviates from the normal phase. ("$b$" stands for "bad.") We define a grim trigger-strategy profile $\bar{\sigma}$ by specifying for each player $i$ and each period $t$ a history-dependent stage-game action $\bar{\sigma}^t_i(h^t)$ according to

---

26 If $\bar{s}$ is an open-loop strategy profile for the infinitely repeated game, then $\bar{s}$ is a Nash equilibrium of the infinitely repeated game if and only if every period's stage-game action profile $\bar{s}^t$ is a Nash equilibrium of the stage game.

$$\bar{\sigma}_i^t(h^t)=\begin{cases}\bar{a}_i(\omega^t), & \hat{\jmath}^t(h^t)=0,\\ b_i^{\hat{\jmath}^t(h^t)}, & \hat{\jmath}^t(h^t)\in I.\end{cases} \tag{21}$$

This says to begin in period zero by choosing your part of the equilibrium-path correlated action profile $\bar{a}(\omega^t)$. Continue to choose this $\omega^t$-dependent stage-game action as long as no player has ever deviated alone. If some player $j$ ever deviates alone, then switch to your part of the stage-game mixed-action profile $b_i^j$ which punishes player $j$, and play that stage-game mixed-action forever, regardless of any future developments. Player $i$'s strategy is shown in flowchart form in Figure 4.

## Nash grim trigger-strategy folk theorem with minmax threats

Let's discover under what conditions $\bar{\sigma}$ is a Nash equilibrium for sufficiently patient players. If all players conform, $\bar{a}(\omega^t)$ is chosen every period. By construction the expectation with respect to the random variable $\omega$ of $g(\bar{a}(\omega))$ is $v$. Therefore the expected repeated-game payoff to each player $i$ if she conforms to $\bar{\sigma}$, given that the other player are conforming as well, is $v_i$. If instead she deviates in some period, play will switch to the open-loop punishment phase. Given that her opponents are playing open-loop strategies in the punishment phase, player $i$'s repeated-game best response is to choose myopic stage-game best responses in each period. She thus receives $\bar{g}_i(b_{-i}^i)$ in every period after her deviation. In other words her conformity and deviation payoff streams have terminal subsequences $v_i$ and $\bar{g}_i(b_{-i}^i)$, respectively. From Theorem 1 we know that a sufficiently patient player will weakly prefer to conform to her part of $\bar{\sigma}$ as long as, for all $i\in I$,

$$v_i>\bar{g}_i(b_{-i}^i). \tag{22}$$

For a given set of $n$ punishment vectors $b^1,\ldots,b^n$, inequalities (22) fully characterize the set of payoffs which can be sustained as Nash equilibria by grim trigger strategies of the form (21). How can we choose the punishment vectors $b^i$ so as to make this equilibrium payoff set as large as possible? Clearly we can do this by making the right-hand side of each inequality (22) as unrestrictive as possible, which we accomplish by minimizing it through our choice of punishment strategy profile. We observe from (17) that $m_{-i}^i$ is exactly the deleted action-profile which minimizes the right-hand side of (22). Therefore by choosing $b^i=m^i$ for each $i\in I$, we have $\bar{g}_i(b_{-i}^i)=\bar{g}_i(m_{-i}^i)=\underline{v}_i$ and therefore establish that every payoff vector $v\gg\underline{v}$ can be sustained as a Nash equilibrium of the repeated game for sufficiently patient players. This is expressed in the following theorem.

**Theorem 4**

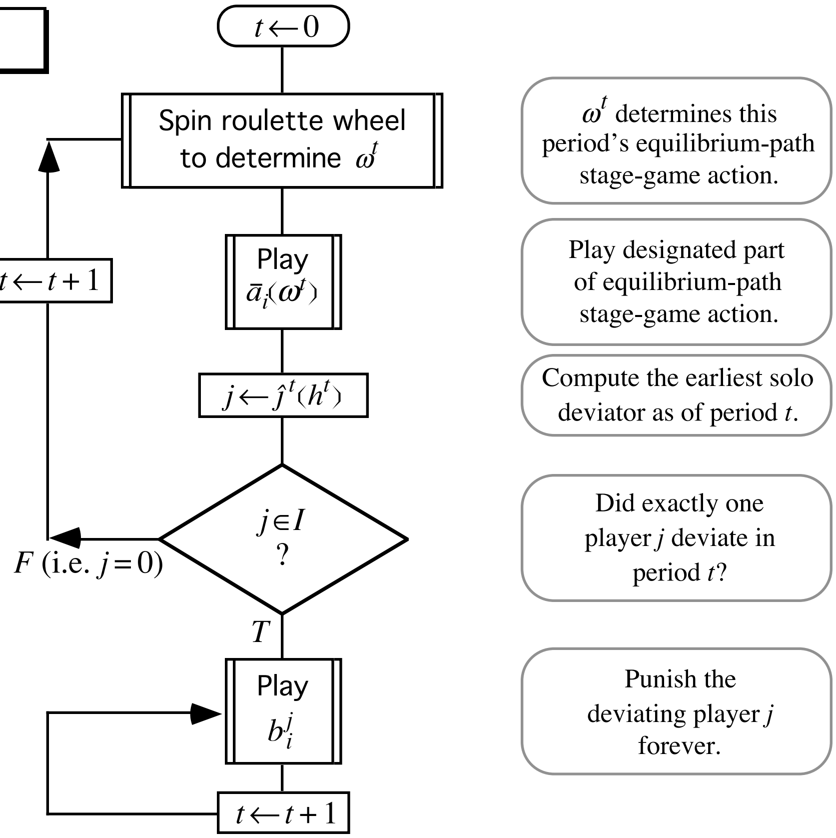


Figure 4: A grim trigger strategy for player $i$.

For any feasible and strictly individually rational payoff vector $v \in V^*$, there exists a repeated-game strategy profile which is a Nash equilibrium of the repeated game for sufficiently patient players and yields the expected payoff vector $v$. I.e., $\forall v \in V^*$, $\exists \bar{\sigma} \in \Sigma$ such that $\exists \underline{\delta} \in (0,1)$ such that $\forall \delta \in (\underline{\delta}, 1)$, $\bar{\sigma}$ is a Nash equilibrium of $G^\infty(\delta)$, and $u(\bar{\sigma}; \delta) = v$.

---

A remark about the definition of the equilibrium strategies is in order: The equilibrium strategy profile $\bar{\sigma}$ specified punishments only when exactly one player deviated from the equilibrium prescription. Why isn't punishment called for if two or more players deviate? When assessing whether she would deviate from her part of an alleged equilibrium strategy profile, each player $i$ asks whether deviation would be profitable given that all other players faithfully fulfill their part of the profile. (This is straight from the definition of Nash equilibrium.) In other words each player $i$'s calculations concern only unilateral deviations by player $i$ herself. The prescriptions in cases of multilateral deviations are of no consequence; we could have assigned any actions in these cases. So why did we expend the effort to define the earliest solo deviator concept? The punishments we specified were tuned to individual players. Unless there was a solo deviator, the target of the punishment would have been ambiguous. In order to use open-loop punishments, we could not allow the target of the punishment to change in response to later player actions; therefore we employed the earliest solo deviator criterion, which has the property that the identity of the punished player never changes.

## Perfect grim trigger-strategy folk theorem with Nash threats

The weakness of Theorem 4 is that it said nothing about whether any feasible and individually rational payoffs are subgame-perfect equilibrium payoffs. We now address the perfection question, while staying within the context of grim trigger strategies. Under what conditions will $\bar{\sigma}$ be a subgame-perfect equilibrium? There are two types of subgames: A normal-phase subgames in which no player has ever deviated solo from the equilibrium path and B punishment-phase subgames in which an open-loop punishment is being played. We already established that no sufficiently patient player would deviate from the normal phase as long as (22) were satisfied for all players. Consider the subgames in class B. These can be further broken down into $n$ groups on the basis of which $b^j$ profile is being played. When all players choose open-loop strategies, the resulting repeated-game strategy profile is a Nash equilibrium of the repeated game if and only if the stage-game action profile played in every period is a Nash equilibrium of the stage game. Therefore, in order that every punishment phase be a Nash equilibrium of that subgame, each $b^i$ must be a Nash-equilibrium action profile of the stage game. So, in summary, $\bar{\sigma}$ is a subgame-perfect equilibrium if, for all $i \in I$,

$$v_i > \bar{g}_i(b^i_{-i}) \quad \text{and} \quad b^i \text{ is a Nash equilibrium of the stage-game.} \tag{23}$$

Let $\mathcal{N} \subset \mathcal{A}$ be the set of stage-game mixed-action Nash equilibria. As we did above with respect to Nash equilibria of the repeated game, we now seek the largest set of payoffs which can be sustained as subgame-perfect equilibria for sufficiently patient players using grim trigger strategies of the form in (21). Again we want to minimize for each player $i$ the right-hand side of inequality (22) through our choice of punishment profile, but now we are restrained to choose each $b^i$ from the Nash equilibrium set $\mathcal{N}$. For each player $i$, we say that a stage-game Nash equilibrium $\eta^i \in \mathcal{N}$ is worst-for-$i$ if

$$\eta^i \in \operatorname*{arg\,min}_{\eta \in \mathcal{N}} g_i(\eta). \tag{24}$$

For any Nash equilibrium $\eta$, $g_i(\eta) = \bar{g}_i(\eta_{-i})$; therefore $\eta^i{}_{-i}$ is the minimizer of $\bar{g}_i(b^i_{-i})$. So choosing, for each $i \in I$, $b^i = \eta^i$, we obtain the following folk theorem.

**Theorem 5** For each player $i \in I$, let $\eta^i$ be a stage-game mixed-action Nash equilibrium which is worst-for-$i$. Define the payoff vector $\tilde{v}$ by $\tilde{v} = (g_1(\eta^1), \ldots, g_n(\eta^n))$. For any feasible and strictly individually rational payoff vector $v$ such that $v \gg \tilde{v}$, there exists a repeated-game strategy profile which is a subgame-perfect equilibrium of the repeated game for sufficiently patient players and yields the expected payoff vector $v$. I.e., $\forall v \in \{v' \in V : v' \gg \tilde{v}\}$, $\exists \bar{\sigma} \in \Sigma$ such that $\exists \underline{\delta} \in (0,1)$ such that $\forall \delta \in (\underline{\delta}, 1)$, $\bar{\sigma}$ is a subgame-perfect equilibrium of $G^\infty(\delta)$ and $u(\bar{\sigma}; \delta) = v$.

---

We know from Theorem 2 that any stage-game Nash equilibrium must give player $i$ at least her minmax value; therefore, for all $i$, $\tilde{v}_i \geq \underline{v}_i$, and therefore $\tilde{v} \geq \underline{v}$.[27] Therefore the set of payoffs supported as subgame-perfect equilibria for sufficiently patient players by Theorem 5 is weakly smaller than the set $V^*$ supported as Nash equilibria by Theorem 4. This leaves open the question of whether the remainder

[27] Recall that, for vectors $x, y \in \mathbb{R}^k$, $x \geq y$ means that, for all $i \in \{1, \ldots, k\}$, $x_i \geq y_i$.

of $V^*$ can be supported in subgame-perfect equilibrium. This will be answered in the affirmative in our last folk theorem.

A weaker version of Theorem 5 was proved by Friedman [1971]. He showed that any payoff vector $v \in V^*$ which strictly Pareto dominates a Nash-equilibrium payoff vector can be supported as a subgame-perfect equilibrium.[28] His punishment phase was the open-loop repetition of a stage-game Nash equilibrium which was dominated by $v$. In other words, he used the same Nash equilibrium to punish any deviator.

For any stage-game Nash-equilibrium action profile $\eta$, $\tilde{v} \le g(\eta)$. There are games such that, for all stage-game Nash equilibria $\eta \in \mathcal{N}$, there is some player $i \in I$ such that $\tilde{v}_i < g_i(\eta)$; in other words, the set of payoffs supported in Theorem 5 is a strictly larger set than that supported by Friedman's theorem. Theorem 5 was able to support the larger set because its punishment profiles were chosen specifically according to the identity of the deviator rather than being "one punishment fits all."

## The ultimate perfect folk theorem

We will now show, subject to a technical qualification concerning the dimensionality of the feasible payoff set $V$, that any feasible and strictly individually rational payoff $v \in V^*$ can be supported as a subgame-perfect equilibrium of an infinitely repeated game with discounting for sufficiently patient players.

Here is the challenge we face…. In order to enforce some payoff vector $v \in V^*$ as an equilibrium, we must be able to punish any deviating player $i$ with a sequence of actions which yields her an average discounted payoff strictly worse than $v_i$. (Otherwise, she would deviate and gladly accept the punishment, since the punishment alone is as least as good as what she would receive in equilibrium.)

Consider a feasible and strictly individually rational payoff vector $v \in V^*$ which we can*not* support as a subgame-perfect equilibrium for sufficiently patient players through the grim trigger strategies with Nash-equilibrium threats of Theorem 5. In other words, there is some player $i \in I$ for whom the payoff $v_i$ she would receive in the specified payoff vector is weakly less than the payoff she would receive in any stage-game Nash-equilibrium. Therefore there is no way that we could punish player $i$ with a stage-game Nash equilibrium so that she would receive a punishment payoff strictly less than $v_i$. Therefore we cannot support this $v \in V$ as a subgame-perfect equilibrium average payoff in the repeated game using open-loop repetitions of stage-game Nash equilibria for punishments. Therefore we cannot support $v$ using grim trigger strategies.

The perfect folk theorem of Fudenberg and Maskin [1986], which we discuss here, solves the problem by punishing a deviator with her minmax profile for only a finite number of periods.

**Theorem 6**

Let $\dim V = n$. For any feasible and strictly individually rational payoff vector $v$, there exists a repeated-game strategy profile which is a subgame-perfect

28 He actually claimed only that they were Nash-equilibrium payoffs. However, his strategies were subgame perfect, so the payoffs are too.

equilibrium of the repeated game for sufficiently patient players and yields the expected payoff vector $v$. I.e., $\forall v\in V^*$, $\exists\bar{\sigma}\in\Sigma$ such that $\exists\underline{\delta}\in(0,1)$ such that $\forall\delta\in(\underline{\delta},1)$, $\bar{\sigma}$ is a subgame-perfect equilibrium of $G^\infty(\delta)$, and $u(\bar{\sigma};\delta)=v$.

---

**Proof** Besides the payoff vector $v$ to be supported and the minmax vector $\underline{v}$, which has already been defined, we need to identify $n+1$ other payoff vectors in order to define the equilibrium strategies. We pick a $\check{v}\in V^*$ which is strictly "between" the minmax vector $\underline{v}$ and the equilibrium payoff $v$, i.e. $\underline{v}\ll\check{v}\ll v$, or equivalently $\forall i\in I$,

$$\underline{v}_i<\check{v}_i<v_i. \qquad (\blacklozenge.1)$$

Now fix some $\varepsilon>0$, and for each player $i\in I$ define the vector $\check{v}^i\in\mathbb{R}^n$ by

$$\check{v}^i_j=\begin{cases}\check{v}_j, & j=i,\\ \check{v}_j+\varepsilon, & j\neq i,\end{cases} \qquad (\blacklozenge.2)$$

for all $j\in I$. In other words, $\check{v}^i$ is a payoff vector which is $\varepsilon$ better than $\check{v}$ for every player except player $i$. See Figure 5. We further require that every such $\check{v}^i$ be a feasible vector; i.e. $\forall i\in I$, $\check{v}^i\in V$. (The existence of such a set of vectors for some $\check{v}$ satisfying ($\blacklozenge$.1) and for some $\varepsilon>0$ is guaranteed by the full-dimensionality assumption.)

In order to keep this presentation of the proof as simple as possible, we'll assume that 1 there is a pure action $\bar{a}\in A$ which yields $v$, i.e. $g(\bar{a})=v$, 2 each player $i$'s minmax profile $m^i$ is a pure-action profile, and 3 each of the $\check{v}^i$ can be achieved via some pure action $\check{a}^i\in A$, i.e. $g(\check{a}^i)=\check{v}^i$.

## The proposed strategies

The proposed equilibrium strategies $\bar{s}_i$, $i\in I$, can best be thought of as specifying a particular stage-game action for player $i$ as a function of what *phase* the game is in. (The phase will be a function of the history; therefore the phase-dependent strategies are history-dependent as well.) The game begins, and in equilibrium remains, in the *normal* phase $N$. In the normal phase players play the action profile $\bar{a}$ each period, which results in the equilibrium payoff vector $v$. If a player $j$ ever deviates solo from her equilibrium-strategy prescription at any point, play switches to the *punish-j* phase $P^j$. This consists of some number $\tau$ of periods of playing the minmax profile $m^j$ in order to punish player $j$ for deviating. (The value of $\tau$ will be specified later.) If this phase ends successfully (no player deviates solo from playing her part of $m^j$ during the $\tau$ periods), the game switches to the *reprieve-j* phase $R^j$. In the $R^j$ phase the action profile $\check{a}^j$ is played each period. Note that play would switch to phase $P^j$ if player $j$ were the sole deviator in any phase—whether she defected from her prescription in the normal phase, in the punishment of another during $P^i$, $i\neq j$, in her own punishment during $P^j$, or during a reprieve phase $R^i$ or $R^j$.

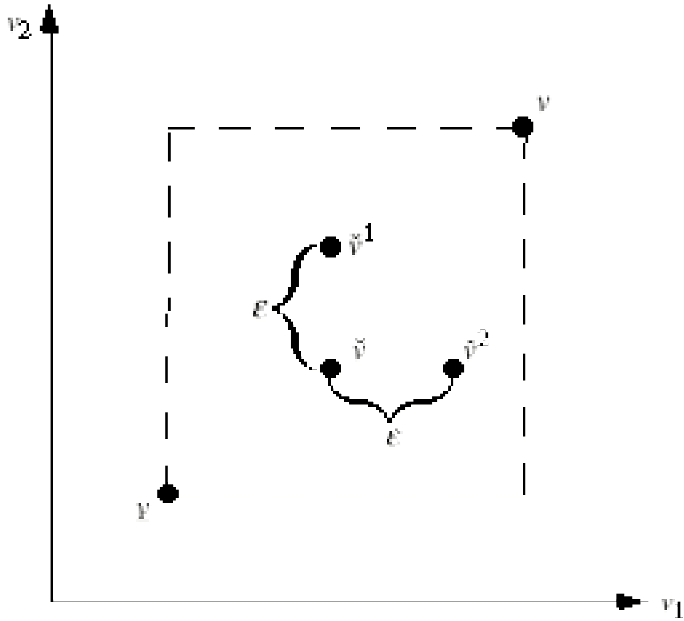


Figure 5: The minmax, equilibrium, and reprieve-phase payoffs.

We let $\varphi^t$ be the phase at the beginning of period $t$, i.e. before the players choose their period-$t$ action. We propose the strategy profile $\bar{s}=(\bar{s}_1,\ldots,\bar{s}_n)$, where each player $i$'s period-$t$ phase-dependent stage-game action is defined by

$$\bar{s}_i^t(\varphi^t)=\begin{cases}\bar{a}_i, & \varphi^t=N,\\ m_i^j, & \varphi^t=P^j,\\ \check{a}_i^j, & \varphi^t=R^j.\end{cases} \tag{◆.3}$$

This rule tells us the action for each phase, but it does not specify how the phase changes during the course of the game.

So we need a phase transition rule. First we define $\tilde{\jmath}(a,a')$ as the solo deviator, if any, between the pair of action profiles $a$ and $a'$. I.e. $\tilde{\jmath}(a,a')=j$ if $a_j\neq a_j{}'$ and $a_{-j}=a_{-j}{}'$. If no player is the solo deviator, then $\tilde{\jmath}(a,a')=0$. We also introduce a *counter*, $\lambda$, for the punishment phases. We set $\varphi^0=N$ and define $\varphi^t$ for $t\in\{1,2,\ldots\}$ by

$$(\varphi^t,\lambda)=\begin{cases}(N,0), & \varphi^{t-1}=N \text{ and } \tilde{\jmath}(a^{t-1},\bar{a})=0,\\ (P^j,1), & \tilde{\jmath}(a^{t-1},\bar{s}^{t-1}(\varphi^{t-1}))=j,\\ (P^j,\lambda+1), & \varphi^{t-1}=P^j,\ \lambda<\tau,\ \text{and } \tilde{\jmath}(a^{t-1},m^j)=0,\\ (R^j,0), & \varphi^{t-1}=P^j,\ \lambda=\tau,\ \text{and } \tilde{\jmath}(a^{t-1},m^j)=0,\\ (R^j,0), & \varphi^{t-1}=R^j \text{ and } \tilde{\jmath}(a^{t-1},\check{a}^j)=0.\end{cases} \tag{◆.4}$$

Let's interpret all this notation a little. The first line says that if the game was in normal phase $N$ in the previous period and no player deviated alone from the normal-phase prescription $\bar{a}$, then play remains in the normal phase. The last line gives similar instructions if the previous period's play had been in a reprieve phase.[29] The second line says that, if in the previous period some player $j$ was the sole deviator from whatever the equilibrium prescription was, viz. $\bar{s}^{t-1}(\varphi^{t-1})$, then switch to the punish-$j$ phase and start the punishment counter at $\lambda = 1$.[30] The third line says that, if play was in a punish-$j$ phase last period and if no player deviated alone from the appropriate punishment profile $m^j$, and as long as there is at least one scheduled punishment unadministered (i.e. $\lambda < \tau$), remain in that punishment phase but increment the punishment counter $\lambda$ by one. The fourth line says that if the last period was the last scheduled punishment of a punishment phase and if that period's punishment was successfully carried out, then switch to the appropriate reprieve phase.

## The easy arguments

Now we show that the strategy profile $\bar{s}$ defined by (◆.3) and the phase-transition rule (◆.4) is a subgame-perfect equilibrium. We invoke the one-stage deviation principle: we check whether there is any player $i$ who can profit by deviating from $\bar{s}_i$ at a single period $t$ and history $h^t$ and then returning to her equilibrium strategy $\bar{s}_i$. If there are no such $i$, $t$, and $h^t$, $\bar{s}$ is a subgame-perfect equilibrium.

Consider what happens when some player $i$ deviates from her equilibrium prescription in some subgame. Play immediately switches to the punishment phase $P^i$. Because all the other players are assumed to be choosing their equilibrium strategies and because we are looking only at one-stage deviations for player $i$, the remainder of the play will be according to $\bar{s}$. We can therefore predict that the punishment phase $P^i$ will be completed successfully and the reprieve phase $R^i$ will be reached and continue forever. This reprieve phase yields the players the payoff vector $\check{v}^i$ in every period. Therefore any player $i$ contemplating a single-stage deviation will face an infinite payoff sequence with a terminal subsequence of $\check{v}_i^i = \check{v}_i$.

For many of this game's subgames we will be able to establish that a sufficiently patient player $i$ would not wish to deviate from $\bar{s}_i$ by the following technique: We will show that conformity leads to an infinite payoff sequence with a terminal subsequence whose value is strictly higher than $\check{v}_i$. Then Theorem 1 tells us that a sufficiently patient player will prefer the conformity payoff stream.

For example, consider any subgame in the normal phase. Conformity to $\bar{s}$ implies that the game will remain in phase $N$ forever, which yields player $i$ a constant payoff stream of $v_i$. (This trivially has a terminal subsequence of $v_i$.) By construction, see (◆.1), $v_i > \check{v}_i$, so no player $i$ will wish to deviate from the normal phase at any point. (I.e. we have shown that no player would deviate at any on-the-equilibrium-path subgame.)

Similarly, consider any subgame in the reprieve phase $R^j$, where $j \neq i$. Conformity would result in player $i$ receiving a payoff of $\check{v}_i^j = \check{v}_i + \varepsilon$ forever. As before, deviation would result in an infinite sequence

---

29 The counter ¬ is not used in either the Normal phase or any reprieve phase, so setting it to zero is an arbitrary decision. The point was just to have the ordered pair $(\varphi^t, \lambda)$ completely defined.

30 The second argument of every occurrence of the $\bar{j}(\cdot,\cdot)$ function could be replaced by $\bar{s}^{t-1}(\varphi^{t-1})$. I use this notation in the second line so that a deviation will be caught, whatever the correct action profile for the phase is. In the other occurrences, on the other hand, I more specifically state what $\bar{s}^{t-1}(\varphi^{t-1})$ should be for the previous period's phase. This is intended to aid your comprehension, not to puzzle you further.

with a terminal subsequence of $\check{v}_i$. Clearly the conformity terminal subsequence has a strictly higher value than that offered by deviation. Therefore no player $i$ would choose to deviate during any reprieve-$j$, $j \neq i$, phase.

The same argument can be made if the game is in the punishment phase $P^j$, $j \neq i$. Conformity by player $i$ implies that the punishment phase will be successfully completed and then the game will switch to the reprieve phase $R^j$, where it will remain. Therefore conformity will also result in player $i$ receiving an infinite payoff sequence with a terminal subsequence of $\check{v}_i + \varepsilon$, which strictly exceeds the $\check{v}_i$ offered by deviation.

So we have shown that a sufficiently patient player $i$ would never deviate from a subgame in phase $N$, $R^j$, or $P^j$, when $j \neq i$. However, we still need to investigate whether player $i$ would deviate from a subgame in phase $R^i$ or $P^i$. In other words, would player $i$ participate in her own reprieve and punishment? We cannot use the previous technique for the following reason. Assume the game is in phase $R^i$ or $P^i$. If player $i$ conforms she faces an infinite stream of payoffs with a terminal subsequence of $\check{v}_i$. As we saw above, if she deviates in any subgame she also faces a terminal subsequence of $\check{v}_i$. So whether she conforms or deviates, her terminal subsequence is the same. Therefore Theorem 1 is agnostic with respect to whether conformity or deviance is preferable.

It is particularly easy to see that player $i$ would not deviate from her own punishment phase $P^i$, thanks to the way the punishment profiles are constructed. When player $i$'s opponents are minmaxing her with $m^i_{-i}$, $\bar{s}_i$ instructs player $i$ to choose her best response $m^i_i$. So player $i$ would gain nothing in the deviation period—because she is already playing a best response—and in fact she would merely prolong the duration of her punishment and postpone her reprieve. To see this more explicitly, assume the game is at a subgame in which there are $\tau' \leq \tau$ periods remaining in the punishment phase. If player $i$ conforms she will earn a continuation payoff of

$$(1-\delta^{\tau'})\underline{v}_i + \delta^{\tau'}\check{v}_i. \qquad (\blacklozenge.5)$$

[See (4).] If she deviates this period, she will earn at most $\underline{v}_i$ this period (since her opponents are minmaxing her), she will earn $\underline{v}_i$ for the next $\tau$ periods of her renewed punishment phase, and then earn $\check{v}_i$ forever. This continuation payoff is at most

$$(1-\delta^{\tau+1})\underline{v}_i + \delta^{\tau+1}\check{v}_i. \qquad (\blacklozenge.6)$$

Since $\tau' < \tau + 1$, the deviation convex combination (◆.6) puts more weight on the strictly lower, minmax payoff than does the conformity payoff (◆.5), therefore it is strictly lower. Therefore no player $i$ would deviate from her own punishment phase $P^i$.[31] This argument is shown more graphically in Figure 6. There we see that the conformity and deviation payoff streams are identical except for $(\tau+1)-\tau'$ periods in which the conformity stage-game payoff is strictly larger than the deviation stage-game payoffs. Therefore the conformity stream dominates the deviation stream for all discount factors.

31 Note that this conclusion required no sufficient-patience limit argument.

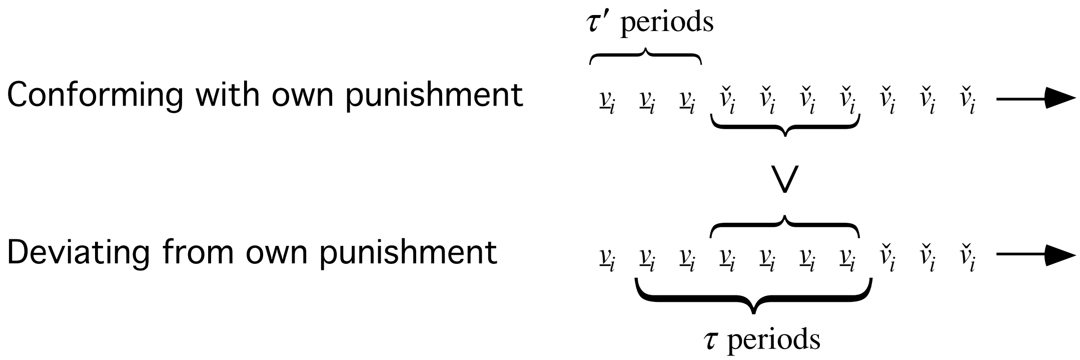


Figure 6: A player would never deviate from her own punishment because the initial deviation gains her nothing and it only prolongs her punishment.

## The harder argument

The only subgames we still have not checked for vulnerability to defection by player $i$ are those in player $i$'s reprieve phase $R^i$. Up to this point we have not needed to specify the duration $\tau$ of the punishment phases; now we will.

If player $i$ conforms with $R^i$, she receives an infinite string of payoffs of $\check{v}_i$. If she deviates in a single stage and conforms thereafter, she receives some unspecified payoff in that period, followed by $\tau$ periods of her minmax payoff from the punishment phase, and then a terminal subsequence of $\check{v}_i$. After the first $\tau+1$ periods (including the deviation period), the two streams are the same: $\check{v}_i$ forever. To determine which stream is preferable, then, we compare the finite discounted sum over the first $\tau+1$ periods.

During the reprieve-$i$ phase the other players are choosing the deleted action profile $\check{a}^i{}_{-i}$. The unspecified deviation-period payoff is certainly bounded above, using (16), by

$$\bar{v}_i \equiv \bar{g}_i(\check{a}^i_{-i}). \tag{◆.7}$$

Using (3), we calculate the discounted sum of the payoffs during these $\tau+1$ periods for both deviation and conformity and require that conformity be weakly preferable. This results in the inequality

$$\check{v}_i + \frac{\delta(1-\delta^\tau)}{1-\delta}\check{v}_i \geq \bar{v}_i + \frac{\delta(1-\delta^\tau)}{1-\delta}\underline{v}_i, \tag{◆.8}$$

where the left-hand side is written unnecessarily expansively to make it more readily comparable to the right-hand side. We rewrite this condition as

$$\Psi(\delta,\tau) \equiv \frac{\delta(1-\delta^\tau)}{1-\delta} \geq \frac{\bar{v}_i - \check{v}_i}{\check{v}_i - \underline{v}_i} \equiv \rho > 0. \tag{◆.9}$$

The question we must answer in the affirmative, if conformity is to be weakly preferable for a sufficiently patient player, is whether for arbitrary $\rho > 0$ there exist $\delta$ and $\tau$ such that the left-hand side of (◆.9) exceeds $\rho$. Two limits are illuminating:

$$\lim_{\tau\to\infty} \Psi(\delta,\tau) = \frac{\delta}{1-\delta}, \tag{◆.10}$$

$$\lim_{\delta\to 1} \Psi(\delta,\tau) = \lim_{\delta\to 1} \frac{1-(\tau+1)\delta^{\tau}}{-1} = \tau, \tag{◆.11}$$

where we used l'Hôpital's rule to evaluate the second limit.[32]

Although the left-hand side function $\Psi(\delta,\tau)$ increases with $\tau$ for fixed $\delta$, (◆.10) shows that it is still bounded above by $\delta/(1-\delta)$ which need not exceed $\rho$. Therefore we cannot satisfy (◆.9) for arbitrary $\delta\in(0,1)$ just by taking the punishment duration $\tau$ sufficiently large. Similarly, although, for fixed $\tau$, $\Psi(\delta,\tau)$ increases with the discount rate $\delta$, (◆.11) shows that $\Psi(\delta,\tau)$ is still bounded above by $\tau$, which need not exceed $\rho$. Therefore we cannot satisfy (◆.9) for an arbitrary punishment duration $\tau$ merely by invoking a sufficient-patience argument.

So we need a combination of taking $\delta$ sufficiently close to 1 and $\tau$ sufficiently large. Limit (◆.11) tells us that $\Psi(\delta,\tau)$ can be made arbitrarily close to $\tau$ by taking $\delta$ sufficiently close to one. By choosing $\tau$ strictly greater than $\rho$, we guarantee that (◆.9) is satisfied by a sufficiently patient player $i$. Hence player $i$ would not deviate from her own reprieve phase. ☺

32 Alternatively, note that $(1-\delta^{\tau})/(1-\delta) = 1+\delta+\delta^2+\cdots+\delta^{\tau-1}$.

# Appendix: The No One-Stage Improvement Principle

Let $s_i=(s_i^0, s_i^1, \ldots, s_i^T)\in S_i$ be a repeated-game strategy for player $i\in I$ in the repeated game "ending" in period $T$, where $T$ may be $\infty$, and where each $s_i^t: A^t \to A_i$. Consider a period $t\in \mathbb{T}\equiv\{0, 1, \ldots, T\}$, a history $h^t=(a^0, \ldots, a^{t-1})\in A^t$, and a strategy profile $s\in S\equiv \mathsf{X}_{i\in I} S_i$. The payoff to player $i$ to the profile $s$ conditional upon the history $h^t$ being reached is

$$u_i(s\,|h^t)=\sum_{\tau=0}^{t-1}\delta^\tau g_i(a^\tau)+\sum_{\tau=t}^{T}\delta^\tau g_i(s^\tau(h^\tau)), \tag{1}$$

where the $h^\tau$, for $\tau>t$, are defined recursively by concatenation as

$$h^\tau=(h^{\tau-1};\ s^{\tau-1}(h^{\tau-1})). \tag{2}$$

The *continuation* payoff to player $i$ to the profile $s$ for the subgame determined by the history $h^t$ is

$$\check{u}_i(s\,|h^t)=\sum_{\tau=t}^{T}\delta^{\tau-t} g_i(s^\tau(h^\tau)). \tag{3}$$

We can relate player $i$'s continuation payoff in adjacent periods by observing from (3) that

$$\check{u}_i(s\,|h^{t-1})=\sum_{\tau=t-1}^{T}\delta^{\tau-(t-1)} g_i(s^\tau(h^\tau))=g_i(s^{t-1}(h^{t-1}))+\delta\sum_{\tau=t}^{T}\delta^{\tau-t} g_i(s^\tau(h^\tau)), \tag{4}$$

which is equivalent to

$$\check{u}_i(s\,|h^{t-1})=g_i(s^{t-1}(h^{t-1}))+\delta\check{u}_i(s\,|(h^{t-1};s^{t-1}(h^{t-1}))), \tag{5}$$

where we have used (2).

A repeated-game strategy profile $s$ is a subgame-perfect equilibrium iff for all players $i\in I$, for all periods $t\in\mathbb{T}$, for all histories $h^t\in A^t$, and for all player-$i$ repeated-game strategies $\tilde{s}_i\in S_i$,

$$u_i((\tilde{s}_i, s_{-i})\,|\,h^t)\le u_i(s\,|\,h^t). \tag{6}$$

Clearly, inequality (6) can be replaced by

$$\check{u}_i((\tilde{s}_i, s_{-i})\,|\,h^t)\le \check{u}_i(s\,|\,h^t). \tag{7}$$

We say that the repeated-game strategy for player $i$, $\hat{s}_i$: $A^T\to A_i^{T+1}$,[33] is a *one-stage deviant* of $s_i$ if there exist a $\hat{t}$ and a history $\hat{h}^{\hat{t}}\in A^{\hat{t}}$ such that 1 $\forall t\neq\hat{t}$, $\hat{s}_i^t=s_i^t$, 2 $\forall h^{\hat{t}}\neq\hat{h}^{\hat{t}}$, $\hat{s}_i^{\hat{t}}(h^{\hat{t}})=s_i^{\hat{t}}(h^{\hat{t}})$, and 3 $\hat{s}_i^{\hat{t}}(\hat{h}^{\hat{t}})\neq s_i^{\hat{t}}(\hat{h}^{\hat{t}})$. Let $\hat{S}_i(s_i)$ be the space of all one-stage deviants of $s_i$.

We say that $s$ satisfies the "no one-stage improvement (NOSI)" property if for all $i\in I$, for all one-stage deviants $\hat{s}_i\in\hat{S}_i(s_i)$ of $s_i$, for all $t\in\mathbb{T}$, and for all $h^t\in A^t$, $\hat{s}_i$ is no better than $s_i$ against $s_{-i}$, conditional

33 This notation seems somewhat nonsensical when $T=\infty$!

on reaching the history $h^t$, i.e.

$$\check{u}_i((\hat{s}_i, s_{-i}) \mid h^t) \leq \check{u}_i(s \mid h^t). \tag{8}$$

It's clear that (8) is a weaker condition than (7); therefore if $s$ is subgame perfect, it satisfies NOSI.

We say that the repeated-game strategy for player $i$, $\tilde{s}_i$: $A^T \to A_i^{T+1}$, is a *finite-stage deviant* of $s_i$ if $\tilde{s}_i$ differs from $s_i$ in at most a finite number of stages. We denote by $\tilde{S}_i(s_i)$ the space of finite-stage deviants of $s_i$. Formally, $\tilde{s}_i \in \tilde{S}_i(s_i)$ iff $\tilde{s}_\mathrm{i} \neq s_i$ and there exists a $\overline{T} < \infty$ such that for all $t \in \{\overline{T}+1, \overline{T}+2, \ldots, T\}$, $\tilde{s}_i^t = s_i^t$.[34]

**Lemma** Let $s \in S$ satisfy the no one-stage improvement property and let $t \in \{1, \ldots, T\}$ be such that for all $i \in I$, for all $h^t \in A^t$, and for all player-$i$ repeated-game strategies $\tilde{s}_i \in S_i$,

$$\check{u}_i((\tilde{s}_i, s_{-i}) \mid h^t) \leq \check{u}_i(s \mid h^t). \tag{9}$$

Then for all $i \in I$, for all $h^{t-1} \in A^{t-1}$, and for all $\tilde{s}_i \in S_i$,

$$\check{u}_i((\tilde{s}_i, s_{-i}) \mid h^{t-1}) \leq \check{u}_i(s \mid h^{t-1}). \tag{10}$$

---

**Proof** Assume not. Then there exist a player $i \in I$, a history $\tilde{h}^{t-1} \in A^{t-1}$, and a player-$i$ repeated game strategy $\tilde{s}_i \in S_i$ such that

$$\check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^{t-1}) > \check{u}_i(s \mid \tilde{h}^{t-1}). \tag{③.1}$$

Rewriting the left-hand side of (③.1), using (5), we obtain

$$\check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^{t-1}) = g_i((\tilde{s}_i^{t-1}, s_{-i}^{t-1})(\tilde{h}^{t-1})) + \delta \check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^t), \tag{③.2}$$

where we have defined

$$\tilde{h}^t = \left(\tilde{h}^{t-1}; (\tilde{s}_i^{t-1}, s_{-i}^{t-1})(\tilde{h}^{t-1})\right). \tag{③.3}$$

Focus now on the last term of (③.2). From (9) we have

$$\check{u}_i(s \mid \tilde{h}^t) \geq \check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^t). \tag{③.4}$$

Therefore, using (③.2) and (③.1),

$$g_i((\tilde{s}_i^{t-1}, s_{-i}^{t-1})(\tilde{h}^{t-1})) + \delta \check{u}_i(s \mid \tilde{h}^t) > \check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^{t-1}) > \check{u}_i(s \mid \tilde{h}^{t-1}). \tag{③.5}$$

Let $\hat{s}_i \in \hat{S}_i(s_i)$ be the one-stage deviant of $s_i$ defined by

$$\hat{s}_i^{t-1}(\tilde{h}^{t-1}) = \tilde{s}_i^{t-1}(\tilde{h}^{t-1}). \tag{③.6}$$

---

34 For a finitely repeated game, just take $\overline{T} = T$.

We observe from (5) that player $i$'s continuation payoff to $\hat{s}_i$ in the subgame determined by $\tilde{h}^{t-1}$, $\check{u}_i((\hat{s}_i, s_{-i}) \mid \tilde{h}^{t-1})$, is exactly the left-hand side of (③.5), because $s_i$ and $\hat{s}_i$ agree for time periods from $t$ onward. Therefore from (③.5)

$$\check{u}_i((\hat{s}_i, s_{-i}) \mid \tilde{h}^{t-1}) > \check{u}_i(s \mid \tilde{h}^{t-1}); \tag{③.7}$$

which contradicts that $s$ satisfies NOSI because $\hat{s}_i$ is a one-stage deviant of $s_i$. ☺

**Theorem 1** If the repeated-game strategy profile $s$ satisfies the no one-stage improvement property then, for every player $i$ and in every subgame $h^t$, $s_i$ is as good for player $i$ against $s_{-i}$ as any finite-stage deviant $\tilde{s}_i \in \tilde{S}_i(s_i)$, conditional upon $h^t$ being reached. In other words, if $s \in S$ satisfies NOSI, then $\forall i \in I$, $\forall t \in \mathbb{T}$, $\forall h^t \in A^t$, $\forall \tilde{s}_i \in \tilde{S}_i(s_i)$,

$$u_i((\tilde{s}_i, s_{-i}) \mid h^t) \leq u_i(s \mid h^t). \tag{11}$$

**Proof** Let $T'$ be such that for all $t \in \{T'+1, T'+2, \ldots, T\}$, $\tilde{s}_i^t = s_i^t$. Let $\overline{T} \equiv \min\{T', T\}$.[35] We establish the premise (9) of the Lemma for the last possibly distinct period $\overline{T}$ and use induction to establish (7) for all $t < \overline{T}$. (Inequality (9) is trivially satisfied for $\overline{T} < t \leq T$.)

Let $t = \overline{T}$ and assume that the premise of the Lemma is not satisfied; i.e. for some $i \in I$, history $\tilde{h}^{\overline{T}}$, and finite-stage deviant $\tilde{s}_i \in \tilde{S}_i(s_i)$,

$$\check{u}_i((\tilde{s}_i, s_{-i}) \mid \tilde{h}^{\overline{T}}) > \check{u}_i(s \mid \tilde{h}^{\overline{T}}). \tag{⑨.1}$$

From (3) we see that (⑨.1) is equivalent to

$$g_i((\tilde{s}_i^{\overline{T}}, s_{-i}^{\overline{T}})(\tilde{h}^{\overline{T}})) > g_i(s^{\overline{T}}(\tilde{h}^{\overline{T}})), \tag{⑨.2}$$

because $\tilde{s}_i$ and $s_i$ agree after period $\overline{T}$. Let $\hat{s}_i \in \hat{S}_i(s_i)$ be the one-stage deviant of $s_i$ defined by $\hat{s}_i^{\overline{T}}(\tilde{h}^{\overline{T}}) = \tilde{s}_i^{\overline{T}}(\tilde{h}^{\overline{T}})$. We can now rewrite (⑨.2) as

$$g_i((\hat{s}_i^{\overline{T}}, s_{-i}^{\overline{T}})(\tilde{h}^{\overline{T}})) > g_i(s^{\overline{T}}(\tilde{h}^{\overline{T}})). \tag{⑨.3}$$

Again appealing to (3) we rewrite (⑨.3) as

$$\check{u}_i((\hat{s}_i, s_{-i}) \mid \tilde{h}^{\overline{T}}) > \check{u}_i(s \mid \tilde{h}^{\overline{T}}), \tag{⑨.4}$$

because $\hat{s}_i$ and $s_i$ agree after period $\overline{T}$. However, (⑨.4) would violate NOSI. Therefore the premise of the Lemma must be satisfied for $t = \overline{T}$.

Now, inductive use of the Lemma establishes (9) for all $t \in \mathbb{T}$ such $t \leq \overline{T}$ and therefore for all $t \in \mathbb{T}$. This condition is exactly (11). ☺

[35] We did not require that $T'$ was the earliest such period. In particular, in a finitely repeated game, it could be the case that $T' > T$.

**Corollary 1** Let $s$ be a strategy profile for a finitely repeated game. The profile $s$ is a subgame-perfect equilibrium of the repeated game if and only if $s$ satisfies the no one-stage improvement property.

---

**Proof** As we observed above, satisfaction of NOSI is necessary for $s$ to be a subgame-perfect equilibrium. In a finitely repeated game, every player-$i$ strategy $\tilde{s}_i \neq s_i$ is a finite-stage deviant of $s_i$. Therefore condition (11) is exactly condition (7) for subgame perfection. ☺

---

Now consider an infinitely repeated game and let $h^\infty$ and $\tilde{h}^\infty$ be any two infinite histories. Let $h^t$ and $\tilde{h}^t$ be their respective restrictions to the first $t$ periods. A game is *continuous at infinity* if

$$\lim_{t\to\infty} \sup_{\substack{h^\infty,\tilde{h}^\infty \in A^\infty \\ h^t = \tilde{h}^t}} \left| u_i(h^\infty) - u_i(\tilde{h}^\infty) \right|. \tag{12}$$

It is easily verified that an infinitely repeated game with discounting is continuous at infinity.

**Theorem 2** Let $s$ be a strategy profile for a finitely repeated game or an infinitely repeated game with discounting. Then $s$ is a subgame-perfect equilibrium if and only if $s$ satisfies the no one-stage improvement property.

---

**Proof** As we observed above, satisfaction of NOSI is necessary for $s$ to be a subgame-perfect equilibrium. Corollary 1 proved sufficiency of NOSI for subgame perfection in the case of finitely repeated games.

Consider an infinitely repeated game with discounting. Assume that $s$ satisfies NOSI but is not a subgame perfect equilibrium. Then there exists a player $i$, a period $\hat{t}$, a history $h^{\tilde{t}}$, and a player-$i$ repeated-game strategy $\hat{s}_i$ such that, for some $\varepsilon > 0$,

$$u_i((\hat{s}_i, s_{-i}) \mid h^{\tilde{t}}) - u_i(s \mid h^{\tilde{t}}) = \varepsilon > 0. \tag{♥.1}$$

(We know that $\hat{s}_i$ is not a finite-stage deviant of $s_i$; this is ruled out by Theorem 1.)

Define

$$\Sigma_{t_1}^{t_2} = \sum_{\tau=t_1}^{t_2} \delta^\tau \Big( g_i((\hat{s}_i^\tau, s_{-i}) \mid h^{\hat{t}}) - g_i(s \mid h^{\hat{t}}) \Big). \tag{♥.2}$$

From (♥.1), we have

$$\varepsilon = \Sigma_{\hat{t}}^{\infty} = \Sigma_{\hat{t}}^{\bar{t}-1} + \Sigma_{\bar{t}}^{\infty}, \tag{♥.3}$$

for any $\bar{t} \in \{\hat{t}, \hat{t}+1, \dots\}$. Because the game is continuous at infinity, we can choose $\bar{t}$ sufficiently large

---

Forcing $\overline{T} \leq T$ ensures that the strategies are defined for $t = \overline{T}$.

that $\Sigma_{\bar{t}}^{\infty} < \frac{1}{2}\varepsilon$. Therefore $\Sigma_{\hat{t}}^{\bar{t}-1} > \frac{1}{2}\varepsilon$.

Now construct a finite-stage deviant of $s_i$ by defining

$$\tilde{s}_i^t = \begin{cases} \hat{s}_i^t, & t < \bar{t}, \\ s_i^t, & t \geq \bar{t}. \end{cases} \quad (\heartsuit.4)$$

We observe that

$$u_i((\tilde{s}_i, s_{-i}) \mid h^{\hat{t}}) - u_i(s \mid h^{\hat{t}}) = \Sigma_{\hat{t}}^{\bar{t}-1} > 0, \quad (\heartsuit.5)$$

because $\tilde{s}_i$ agrees with $\hat{s}_i$ prior to period $\bar{t}$ and agrees with $s_i$ from $\bar{t}$ onward. Because (♥.5) is positive, the conclusion of Theorem 1 is violated and therefore, contrary to assumption, $s$ must not satisfy NOSI. ☺

# References

Aumann, Robert J. [1985] "Repeated Games," in *Issues in Contemporary Microeconomics and Welfare*, ed. George Fiewel, MacMillan, pp. 209–242.

Benoît, Jean-Pierre and Vijay Krishna [1985] "Finitely Repeated Games," *Econometrica* **53** 4 (July): 905–922.

Friedman, James W. [1971] "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies* **38** 113 (January): 1–12.

Fudenberg, Drew and Eric S. Maskin [1986] "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica* **54** 3 (May): 533–554.

Fudenberg, Drew and Jean Tirole [1991] *Game Theory*, MIT Press.

Intriligator, Michael D. [1971] *Mathematical Optimization and Economic Theory*, Series: Prentice-Hall Series in Mathematical Economics, Prentice-Hall, Inc.

Myerson, Roger B. [1991] *Game Theory: Analysis of Conflict*, Harvard University Press.

Nalebuff, Barry [1988] "Puzzles: Cider in Your Ear, Continuing Dilemma, The Last Shall Be First, and More," *Journal of Economic Perspectives* **2** 2 (Spring): 149–156.

Rubinstein, Ariel [1979] "Equilibrium in Supergames with the Overtaking Criterion," *Journal of Economic Theory* **21**: 1–9.

Sorin, Sylvain [1986] "On Repeated Games with Complete Information," *Mathematics of Operations Research* **11** 1 (February): 147–160.