Infinitely Repeated Games

Recall the following theorem

Theorem 72 *If a game has a unique Nash equilibrium, then its finite repetition has a unique SPNE.*

Our intuition, however, is that long-term relationships may be fundamentally different from one-shot meetings. This is one of the reasons that we consider infinite repetitions of games. Infinitely repeated games also model a long-term relationship in which the players do not know a priori when they will stop repeating the game: there is no pre-ordained number of repetitions.

Recall the terminology: The game that is being repeated is the **stage game**. The stages of the game are $t = 0, 1, 2, ...$ An infinitely repeated game is also sometimes called a **supergame**.

How players evaluate payoffs in infinitely repeated games. A player receives an infinite number of payoffs in the game corresponding to the infinite number of plays of the stage game. We need a way to calculate a finite payoff from this infinite stream of payoffs in order that a player can compare his strategies in the infinitely repeated game.

There are two alternative approaches. Let ρ_{it} denote the payoff that player i receives in the tth stage of the game.

The most widely used approach is **discounted payoffs**. Let δ_i denote player i's discount factor. Player i evaluates the infinite sequence of payoffs

$$\rho_{i0}, \rho_{i1}, \rho_{i2}, \cdots$$

as the sum

$$\sum_{t=0}^{\infty} \delta_i^t \rho_{it}.$$

This will be a finite number as long as $(|\rho_{it}|)_{t=0}^{\infty}$ is bounded above. This discounted sum is typically modified by putting $(1 - \delta_i)$ in front,

$$(1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t \rho_{it}.$$

This is a renormalization of utility that doesn't change player i's ranking of any two infinite sequences of payoffs. The $(1 - \delta_i)$ insures that player i evaluates the sequence in which he receives a constant c in each period as c, i.e.,

$$
\begin{aligned}
(1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t c &= c(1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t \\
&= c(1 - \delta_i) \frac{1}{(1 - \delta_i)} \\
&= c,
\end{aligned}
$$

where we have applied the formula for the sum of a geometric series.

An alternative approach is **limiting average payoffs**. It is sometimes simpler to use than discounted payoffs. This leaves open the question, "Which formula is the best model for a person's preferences over time?"

Player i evaluates the infinite sequence of payoffs

$$\rho_{i0}, \rho_{i1}, \rho_{i2}, \cdots$$

as the limit

$$\lim_{n \to \infty} \frac{\sum_{t=0}^{n-1} \rho_{it}}{n}.$$

The existence of this limit is sometimes a problem. The advantage of this formula, however, is that it is easy to calculate the limiting payoff if the sequence of payoffs

$$\rho_{i0}, \rho_{i1}, \rho_{i2}, \cdots$$

eventually reaches some constant payoff,

$$\rho_{it} = c \text{ for all } t \geq t'$$

regardless of the payoffs over stages 0 through $t'-1$. We will use limiting average payoffs below in some cases to simplify the analysis of infinitely repeated games. This approach is particularly useful in introducing the Folk Theorem.

Example 73 (An Infinitely Repeated Prisoner's Dilemma) *We'll analyze the game using discounted payoffs. Consider the following version of the prisoner's dilemma:*

1/2	c	nc
c	2,2	-3,3
nc	3,-3	-2,-2

Here, c refers to "cooperate" (not "confess") while nc refers to "don't cooperate". A theorem that we stated in the beginning of the class implies that there is a unique SPNE in the finite repetition of this game, namely nc, nc in each and every stage.

This remains an SPNE outcome of the infinitely repeated game. Consider the strategies:

$$\begin{aligned} 1 \quad &: \quad \textit{play nc in every stage} \\ 2 \quad &: \quad \textit{play nc in every stage.} \end{aligned}$$

Given the other player's strategy, playing nc maximizes player i's payoff in each stage of the game and hence maximizes his discounted payoff (and also his average payoff, if that is how he's calculating his return in the infinite game). This isn't a very interesting equilibrium, however; why bother with infinite repetition if this is all that we can come up with? In particular, we ask "Can the players sustain c,c as the outcome in each and every stage of the game as a noncooperative equilibrium"?

Consider the following strategy as played by both players:

1. *play c to start the game and as long as both players play c;*

2. *if any player ever chooses nc, then switch to nc for the rest of the game.*

*This is a **trigger strategy** in the sense that bad behavior (i.e., playing nc) by either player triggers the punishment of playing nc in the remainder of the game. It is sometimes also called a "grim" trigger strategy to emphasize how unforgiving it is: if either player ever chooses nc, then player i will punish his opponent forever.*

Does the use of this trigger strategy define an SPNE? Playing c in any stage does not maximize a player's payoff in that stage (nc is the best response within a stage). Suppose player i starts with this strategy and considers deviating in stage k to receive a payoff of 3 instead of 2. Thereafter, his opponent chooses nc, and so he will also choose nc in the remainder of the game. The use of trigger strategies therefore defines a Nash equilibrium if and only if the equilibrium payoff of 2 in each stage is at least as large as the payoff from deviating to nc in stage k and ever thereafter:

$$(1-\delta_i)\sum_{t=0}^{\infty}\delta_i^t \cdot 2 \quad \geq \quad (1-\delta_i)\left[\sum_{t=0}^{k-1}\delta_i^t \cdot 2 + \delta_i^k \cdot 3 + \sum_{t=k+1}^{\infty}\delta_i^t \cdot (-2)\right] \Leftrightarrow$$

$$\sum_{t=k}^{\infty}\delta_i^t \cdot 2 \quad \geq \quad \delta_i^k \cdot 3 + \sum_{t=k+1}^{\infty}\delta_i^t \cdot (-2) \Leftrightarrow$$

$$\sum_{t=k}^{\infty}\delta_i^t \cdot 2 \quad \geq \quad \delta_i^k \cdot 3 + \sum_{t=k+1}^{\infty}\delta_i^t \cdot (-2) \quad \textit{cancel the first k terms}$$

$$\Leftrightarrow \quad \sum_{t=0}^{\infty}\delta_i^t \cdot 2 \geq 3 + \sum_{t=1}^{\infty}\delta_i^t \cdot (-2) \quad \textit{cancel } \delta_i^k$$

$$\frac{2}{1-\delta_i} \quad \geq \quad 3 + \frac{\delta_i}{1-\delta_i}(-2)$$

$$2 \quad \geq \quad 3 - 3\delta_i - 2\delta_i$$

$$5\delta_i \quad \geq \quad 1$$

$$\delta_i \quad \geq \quad \frac{1}{5}$$

47

Deviating from the trigger strategy produces a one-time bonus of changing one's stage payoff from 2 to 3. The cost, however, is a lower payoff ever after. We see that the one-time bonus is worthwhile for player i only if his discount factor is low ($\delta_i < 1/5$), so that he doesn't put much weight upon the low payoffs in the future.

When each $\delta_i \geq 1/5$, do the trigger strategies define a subgame perfect Nash equilibrium (in addition to being a Nash equilibrium)? Yes.

*A subgame of the infinitely repeated game is determined by a **history**, or a finite sequence of plays of the game. There are two kinds of histories to consider:*

1. *If each player chose c in each stage of the history, then the trigger strategies remain in effect and define a Nash equilibrium in the subgame.*

2. *If some player has chosen nc in the history, then the two players use the strategies*

$$1 \quad : \quad \text{play nc in every stage}$$
$$2 \quad : \quad \text{play nc in every stage.}$$

in the subgame. As we discussed above, this is a Nash equilibrium.

Therefore, whichever of the two kinds of history we have, the strategies define a Nash equilibrium in the subgame. The trigger strategies therefore define a subgame perfect Nash equilibrium whenever they define a Nash equilibrium.

Recall the fundamental importance of the Prisoner's Dilemma: it illustrates quite simply the contrast between self-interested behavior and mutually beneficial behavior. The play of nc,nc instead of c,c represents the cost of noncooperative behavior in comparison to what the two players can achieve if they instead were able to cooperate. What we've shown is that that the cooperative outcome can be sustained as a noncooperative equilibrium in a long-term relationship provided that the players care enough about future payoffs.

A General Analysis

We let (s_1, s_2) denote a Nash equilibrium of the stage game with corresponding payoffs (π_1, π_2). Suppose that the choice of strategies (s_1^*, s_2^*) would produce the payoffs (π_1^*, π_2^*) where

$$\pi_i^* > \pi_i$$

for each player i. The strategies (s_1^*, s_2^*) would therefore produce a better outcome for each player. The strategies (s_1^*, s_2^*) are not a Nash equilibrium, however; when player $-i$ chooses s_{-i}^*, the maximal payoff that player i can achieve by changing his strategy away from s_i^* is $d_i > \pi_i^*$. Note that we are assuming that

$$d_i > \pi_i^* > \pi_i. \tag{1}$$

Can trigger strategies sustain the use of the strategies (s_1^*, s_2^*) in each and every stage of the game? The trigger strategy here for each player i is:

1. play s_i^* to start the game and as long as both players play (s_1^*, s_2^*);

2. if any player ever deviates from the pair (s_1^*, s_2^*) then switch to s_i for every stage in the remainder of the game.

We'll calculate a lower bound on δ_i that is sufficient to insure that player i will not deviate from s_i^*. Suppose player i deviates from s_i^* in stage k. We make two observations:

1. Player $-i$ switches to s_{-i} in each stage $t > k$. Player i's best response is to choose s_i in each stage after the kth (recall our assumption that (s_1, s_2) is a Nash equilibrium).

2. The maximal payoff that player i can gain in the kth stage is d_i (by assumption).

The following inequality is therefore necessary and sufficient for player i to prefer his trigger strategy to the deviation that we are considering:

$$(1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t \cdot \pi_i^* \geq (1 - \delta_i) \left[\sum_{t=0}^{k-1} \delta_i^t \cdot \pi_i^* + \delta_i^k \cdot d_i + \sum_{t=k+1}^{\infty} \delta_i^t \cdot \pi_i \right] \Leftrightarrow$$

$$\sum_{t=k}^{\infty} \delta_i^t \cdot \pi_i^* \geq \delta_i^k \cdot d_i + \sum_{t=k+1}^{\infty} \delta_i^t \cdot \pi_i \Leftrightarrow$$

$$\sum_{t=k}^{\infty} \delta_i^t \cdot \pi_i^* \geq \delta_i^k \cdot d_i + \sum_{t=k+1}^{\infty} \delta_i^t \cdot \pi_i \text{ cancel the first } k \text{ terms}$$

$$\Leftrightarrow \sum_{t=0}^{\infty} \delta_i^t \cdot \pi_i^* \geq d_i + \sum_{t=1}^{\infty} \delta_i^t \cdot \pi_i \text{ cancel } \delta_i^k$$

$$\frac{\pi_i^*}{1 - \delta_i} \geq d_i + \frac{\delta_i}{1 - \delta_i} \cdot \pi_i$$

$$\pi_i^* \geq d_i (1 - \delta_i) + \delta_i \pi_i$$

$$\pi_i^* - d_i \geq \delta_i (\pi_i - d_i)$$

$$d_i - \pi_i^* \leq \delta_i (d_i - \pi_i)$$

$$\delta_i \geq \frac{d_i - \pi_i^*}{d_i - \pi_i} \text{ applying (1)}$$

As in the previous example, we have obtained a lower bound on δ_i that is sufficient to insure that player i will not deviate from his trigger strategy given that the other player uses his trigger strategy.

Several observations are in order:

1. The analysis focuses on a single player at a time and exclusively on his payoffs. The bound thus extends immediately to stage games with $n > 2$ players. The assumption that there are two players has no role in the above analysis.

2. Notice that any player who deviates from the "better" strategies (s_1^*, s_2^*) triggers the switch by both players to the Nash equilibrium strategies (s_1, s_2). This is unfair in the sense that both players suffer from the bad behavior of one of the two players (it is part of the definition of equilibrium).

3. If $d_i \leq \pi_i^*$, then player i has no incentive to deviate from s_i^* (he doesn't even get a one-stage "bonus" from ending the play of (s_1^*, s_2^*) for the rest of the game). We thus don't have to worry about player i's willingness to stick to his trigger strategy regardless of the value of his discount factor.

Example 74 *Consider the following stage game:*

$1 \backslash 2$	L	C	R
T	1,-1	**2,1**	1,0
M	**3,4**	0,1	-3,2
B	4,-5	-1,3	1,1

The unique pure strategy Nash equilibrium is T,C, which gives the payoffs 2,1. Both players prefer the outcome 3,4 determined by the play of M,L, which isn't a Nash equilibrium. We consider the trigger strategies

- *Player 1: Play M to start the game and as long as the strategies M,L are played; if M,L is ever not played, then switch to T for all future stages of the game.*

- *Player 2: Play L to start the game and as long as the strategies M,L are played; if M,L is ever not played, then switch to C for all future stages of the game.*

From above we have the following bound on player 1's discount factor:

$$\delta_1 \geq \frac{d_1 - \pi_1^*}{d_1 - \pi_1} = \frac{4 - 3}{4 - 2} = \frac{1}{2}.$$

If $\delta_1 \geq 1/2$, then player 1's trigger strategy is a best response to player 2's trigger strategy. Given that player 1 plays M, player 2's best response is L (he has no reason to switch to any other strategy). Player 2's trigger strategy is thus a best response to player 1's trigger strategy for all values of δ_2.

An Application in Industrial Organization

One form of collusive behavior among firms is to restrict output in order to keep the price of the product high. This is a goal of the OPEC oil cartel, for instance: member countries have output quotas that are mutually negotiated within the cartel with an eye toward keeping the price of oil at a desired level. Economists have often claimed that cartels and collusive behavior are fundamentally unstable and hence unlikely to endure for the long-term. The argument essentially is that collusive agreements are not Nash equilibria and member firms or countries have the incentive to cheat on the common agreement in pursuit of their own self-interests. This causes the cartel to break down. A footnote to this is that agreements within a nation among firms may violate antitrust laws. The firms therefore have no legal means of contracting among themselves or appealing to the courts for punishment if one or more firms fails to live up to its obligations to the cartel. Similar concerns apply for cartels of nations who have no over-arching government that can enforce the mutually beneficial arrangement.

How then do we explain the fact that OPEC has for the most part succeeded for over 50 years in influencing the global price of oil? Or how do we explain the documented existence of cartels among industries (such as railroads in the U.S. in the late 19th century)? Our model above suggests an answer. A collusive arrangement provides each firm with a larger profit than the competitive outcome. The collusive arrangement is a noncooperative equilibrium in a long-term relationship, provided that each firm cares enough about future profits.

Example 75 (Cournot Duopoly) *We illustrate the point in a simple example. There are two identical firms that produce the same product. Let q_i denote the output of firm i. The market price for the aggregate output $q = q_1 + q_2$ is determined by the inverse demand function*

$$p(q) = 14 - q.$$

The cost function of each firm i is

$$C(q_i) = \frac{q_i^2}{4}.$$

The profit function of firm i is therefore

$$P_i(q_i, q_{-i}) = (14 - (q_1 + q_2))\, q_i - \frac{q_i^2}{4}.$$

The Nash equilibrium outputs. *We solve for a Nash equilibrium by setting $\partial P_i/\partial q_i = 0$ for each firm i:*

$$\frac{\partial P_1}{\partial q_1} = (14 - (q_1 + q_2)) - q_1 - \frac{q_1}{2} = 0,$$

$$\frac{\partial P_2}{\partial q_2} = (14 - (q_1 + q_2)) - q_2 - \frac{q_2}{2} = 0,$$

or

$$14 - \frac{5q_1}{2} - q_2 = 0,$$

$$14 - q_1 - \frac{5q_2}{2} = 0.$$

This implies

$$14 - \frac{5q_1}{2} - q_2 = 14 - q_1 - \frac{5q_2}{2}$$

or $q_1 = q_2$. Substitution into either equation implies

$$14 - \frac{5q_1}{2} - q_1 = 0$$

$$14 = \frac{7}{2}q_1 \Rightarrow$$

$$q_1 = q_2 = 4$$

From above we see that

$$\frac{\partial P_i}{\partial q_i}(q_i, q_{-i} = 4) = (14 - (q_i + 4)) - \frac{3q_i}{2}$$

$$= 10 - \frac{5q_i}{2}$$

which changes from positive to negative at $q_i = 4$. This verifies that $q_1 = q_2 = 4$ is a Nash equilibrium. The Nash equilibrium profit for each firm is

$$(14 - (8))\, 4 - \frac{16}{4} = 24 - 4 = 20.$$

A better outcome for the firms. *We now calculate the outputs q_1, q_2 that maximize the sum of the profits for the two firms:*

$$P_1(q_1, q_2) + P_2(q_1, q_2) = (14 - (q_1 + q_2))\,(q_1 + q_2) - \frac{q_1^2}{4} - \frac{q_2^2}{4}.$$

$$0 = \frac{\partial}{\partial q_1} = (14 - (q_1 + q_2)) - (q_1 + q_2) - \frac{q_1}{2}$$

$$0 = 14 - \frac{5q_1}{2} - 2q_2,$$

$$0 = \frac{\partial}{\partial q_2} = 14 - 2q_1 - \frac{5q_2}{2}$$

Again, we have $q_1 = q_2$. Solving using either partial derivative implies

$$0 = 14 - \frac{9}{2}q_i$$

or $q_1 = q_2 = \frac{28}{9}$.

I don't want to work with such awful fractions. These numbers suggest that both firms would obtain a higher profit than in the Nash equilibrium by choosing $q_1 = q_2 = 3$ (which is close to 28/9). Let's check:

$$(14 - 6)\, 3 - \frac{9}{4} = 24 - \frac{9}{4} = \frac{87}{4} > 20.$$

If firm $-i$ produces 3, then how much profit can firm i obtain by deviating? *Firm i maximizes*

$$P_i(q_i, 3) = (14 - (q_i + 3))\, q_i - \frac{q_i^2}{4}$$

$$= (11 - q_i)\, q_i - \frac{q_i^2}{4}$$

$$= 11q_i - \frac{5q_i^2}{4}$$

$$0 = \frac{\partial P_i}{\partial q_i} = 11 - \frac{5q_i}{2} \Rightarrow q_i = \frac{22}{5}$$

The marginal profit changes sign at $q_i = \frac{22}{5}$, and so it indeed maximizes profit. The maximally profitable deviation therefore produces a profit of

$$P_i\left(\frac{22}{5}, 3\right) = \frac{22}{5}\left(11 - \frac{5}{4} \cdot \frac{22}{5}\right)$$

$$= \frac{22}{5}\left(\frac{11}{2}\right)$$

$$= \frac{121}{5}.$$

Implementation of the superior outcome. *Each firm i adopts the following strategy: Choose $q_i = 3$ to start the game and as long as each firm produces 3 units in each stage game. If any output is observed by either firm other than 3, then switch to $q_i = 4$ for every stage in the remainder of the game.*

The Nash equilibrium profit is

$$\pi_i = 20,$$

a "collusive" outcome produces the profit

$$\pi_i^* = \frac{85}{4},$$

and each firm can deviate from the collusive output of 3 to obtain

$$d_i = \frac{121}{5}$$

The use of this trigger strategy by each firm defines a subgame perfect Nash equilibrium in the infinitely repeated Cournot duopoly game if each firm i's discount factor δ_i satisfies

$$\begin{aligned}
\delta_i \quad &\geq \quad \frac{d_i - \pi_i^*}{d_i - \pi_i} = \frac{\frac{121}{5} - \frac{87}{4}}{\frac{121}{5} - 20} = \frac{\frac{121}{5} - \frac{87}{4}}{\frac{21}{5}} \\
&= \quad \frac{484 - 435}{84} = \frac{49}{84} = \frac{7}{12}.
\end{aligned}$$

Several Observations:

- The relevance of this result in the theory of repeated games to explaining how collusion occurs despite the absence of legal structures to enforce the collusive agreement was first noted by Jim Friedman.

- Are trigger strategies realistic? We in fact see cartels sustaining their collusive behavior through mutual punishment if anyone member cheats on the collusive agreement. As the member of OPEC with the greatest reserves and capacity, Saudi Arabia plays the role of enforcer in the following sense. Suppose some nation cheats by producing beyond its OPEC-negotiated quota. Saudi Arabia opens its taps and floods the world market with oil, punishing all members with a lower price for oil and correspondingly low profits. After a period of punishment, the cartel gets its act together and reinstitutes a collusive agreement. Such flooding of the market has happened several times in the history of OPEC. It is the tool or threat that Saudi Arabia has to keep the member countries in line.

- The preceding story about punishment, however, does not correspond to an equilibrium in trigger strategies. Notice that: (i) in an equilibrium with trigger strategies, the firms collude and never revert to the Nash equilibrium outputs; (ii) if the firms ever did switch to the Nash equilibrium outputs, they would do so forever and would never reestablish the collusive outcome. These issues have been addressed in a paper by Ed Green and Rob Porter.[2] Each firm in this paper observes the market price and not the output of the other firm. Moreover, there is a random or stochastic element to market demand in their model; a decline in the market price may therefore be caused by an increase in production by a firm or simply by a random decline to demand. Green and Porter construct equilibria of the infinitely repeated game in which:

  - Each firm starts out producing at a collusive level;
  - A market price that falls below a target $\widetilde{p}$ causes the two firms to enter a *punishment phase* in which they each choose larger and less profitable outputs for n stages;
  - After the n stages of the punishment phase, each firm returns to its collusive output level.

  The target price $\widetilde{p}$ and the length n of the punishment phase are part of the construction of the equilibrium. Their equilibrium has the property that (i) periods of intense competition through overproduction between the two firms occur with positive probability during the infinitely repeated game, and (ii) after a punishment phase, the firms reestablish their collusive agreement (that is, until it breaks down again). In equilibrium, no firm ever deviates from the collusive output in non-punishment stages; the punishment phases occur with positive probability, however, because of random declines in the market price.

---

[2] Edward J. Green and Robert H. Porter, "Noncooperative Collusion Under Imperfect Price Information", *Econometrica*, Vol. 52 (1984), p. 87-100.

Minmax values

Our definition of a trigger strategy has a player switch to a Nash equilibrium strategy in the event that punishment is triggered in the game. Assuming that the discount factors are sufficiently large so that the trigger strategies form a Nash equilibrium, the assumption that the players switch to a Nash equilibrium if punishment is triggered insures that the equilibrium is subgame perfect, i.e., the punishment is credible. We can obtain a smaller "lower bound" on the discount factor for Nash equilibrium if we dispense with the requirement of subgame perfection. Let's think instead about the worst punishment that one player can impose on the other because this will serve as the most effective deterrent.

Player i's **minmax value** m_i is the lowest payoff in the stage game that player $-i$ can impose on him through his choice of a strategy s_{-i} given that player i can choose his own strategy to maximize his own payoff:

$$m_i = \min_{s_{-i}} \max_{s_i} \pi_i \left(s_i, s_{-i} \right).$$

Here, $\pi_i \left(s_i, s_{-i} \right)$ is the payoff to player i in the stage game given the strategy profile (s_i, s_{-i}). The "max" represents the capability of player i to choose his own strategy s_i to maximize his payoff given the other player's strategy s_{-i}, and the "min" represents player $-i$'s choice of s_{-i} minimize this "best response" payoff for player i. Let $\widetilde{s}_{-i}$ denote the strategy of player $-i$ at which the minmax value is obtained. The strategy $\widetilde{s}_{-i}$ is the worst choice of a strategy by player $-i$ from the perspective of player i. Existence of $\widetilde{s}_{-i}$ is not a problem in finite games.

Notice that player i cannot receive less than his minmax value in any Nash equilibrium of the stage game. The value m_i is the lowest payoff that player i receives when he chooses his strategy in his own best interest.

We can of course define a minmax value m_{-i} for player $-i$ and the strategy $\widetilde{s}_i$. For a desired outcome (s_1^*, s_2^*) in the stage game with corresponding payoffs (π_1^*, π_2^*), define the trigger strategy of player i as follows:

Player i: Choose s_i^* to start the game and as long as (s_1^*, s_2^*) is played by the players; if (s_1^*, s_2^*) is ever not played, then switch to $\widetilde{s}_i$ for ever more.

The choice of $\widetilde{s}_i$ in every future stage is the worst thing that player i can do to player $-i$ and it therefore is the most effective deterrent. Similar remarks hold for player $-i$'s choice of $\widetilde{s}_{-i}$. We obtain a different and looser lower bound on the discount factors in this case. Player i compares his payoff if he follows his trigger strategy to his best possible deviation in stage k (assuming that $d_i > \pi_i^*$):

$$(1 - \delta_i) \sum_{t=0}^{\infty} \delta_i^t \cdot \pi_i^* \quad \geq \quad (1 - \delta_i) \left[\sum_{t=0}^{k-1} \delta_i^t \cdot \pi_i^* + \delta_i^k \cdot d_i + \sum_{t=k+1}^{\infty} \delta_i^t \cdot m_i \right] \Leftrightarrow$$

$$\delta_i \quad \geq \quad \frac{d_i - \pi_i^*}{d_i - m_i}$$

We substitute his minmax payoff for his Nash equilibrium payoff π_i. Notice that

$$\frac{d_i - \pi_i^*}{d_i - m_i} \leq \frac{d_i - \pi_i^*}{d_i - \pi_i}$$

because $m_i \leq \pi_i$ (a player's minmax value is less than or equal to his payoff in any Nash equilibrium).

These new trigger strategies do not necessarily form a subgame perfect Nash equilibrium when they do form a Nash equilibrium. Consider a history in which (s_1^*, s_2^*) does not occur in some stage. In the subgame defined by that history, the strategies specify that the players play $(\widetilde{s}_i, \widetilde{s}_{-i})$ in each and every stage. This need not define a Nash equilibrium in the subgame (in particular, there is no reason that $\widetilde{s}_i$ must be a best response to $\widetilde{s}_{-i}$ in the stage game).

Through the use of a more severe punishment by each player in their trigger strategies, we have obtained a lower bound on the discount factor δ_i that is sufficient to insure that (π_1^*, π_2^*) is played in each and every stage of a Nash equilibrium of the supergame. We have sacrificed subgame perfection of the equilibrium, however, in that punishment may not be credible.

As a final point, note that the analysis would change further if we considered mixed strategies. Such strategies can lower or raise a player's minmax value (the set of strategies over which the min is taken is increased in size, but so is the set of strategies over which the max is taken). This really doesn't change any of the ideas that we are discussing, however, and so we'll stick to the case of pure strategies.

Example 76 *Let's reconsider the following stage game:*

1\2	L	C	R
T	1,-1	2,1	1,0
M	**3,4**	0,1	-3,2
B	4,-5	-1,3	1,1

We'd like to implement $(3,4)$ as the outcome in each stage. As above, we don't need to worry about player 2 deviating from his trigger strategy. Let's focus on player 1's minmax value m_1 and the strategy $\widetilde{s}_2$ that player 2 should use if he really wants to hurt player 1:

$$m_1 = 1, \widetilde{s}_2 = R.$$

Before we had the following bound on player 1's discount factor:

$$\delta_1 \geq \frac{d_1 - \pi_1^*}{d_1 - \pi_1} = \frac{4-3}{4-2} = \frac{1}{2}.$$

*If player 2 punishes with R instead of C, we have the bound*

$$\delta_1 \geq \frac{d_1 - \pi_1^*}{d_1 - m_1} = \frac{4-3}{4-1} = \frac{1}{3}.$$

A Version of the Folk Theorem

We've discussed implementing as a Nash in the supergame any outcome that gives each player a larger per stage payoff than he receives in a Nash equilibrium of the stage game. Let's return to the prisoner's dilemma and depict graphically all of the possible payoffs of the players in the supergame. This discussion will be easier if we now switch to limiting average payoffs as the method used by players to calculate their payoffs in the infinitely repeated game.

| 1/2	c	nc
c	2,2	-3,3
nc	3,-3	-2,-2

The average payoffs of the two players in the supergame is a point in the convex hull of the four pairs $(2,2)$, $(-3,3)$, $(3,-3)$, and $(-2,-2)$:
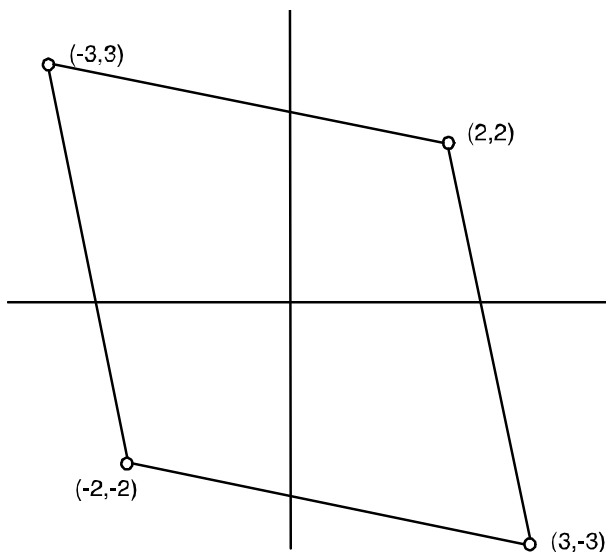
The Folk Theorem describes the points in this convex hull that can result from the play of Nash equilibria in the supergame. The term "Folk" refers to the fact that the result was known in the small community of game theorists in the 1960s before anyone wrote it down or formalized the proof. It was like a "folk" song, whose origin in unknown and which is passed among people by oral communication. It would really be more correct to say "a" Folk Theorem because there are variations that depend upon how payoffs are calculated in the supergame, whether or not a refinement of Nash equilibrium such as subgame perfection

is added, etc.. We'll discuss some of these alternative versions after presenting result in the simple case of average payoffs.

The average payoff in an equilibrium is a weighted average of the four outcomes $(2,2)$, $(-3,3)$, $(3,-3)$, and $(-2,-2)$ where the weights reflect the frequency with which the different outcomes are played in the equilibrium. We note first that each player cannot receive less than his minmax value as his average payoff in the supergame. This is true because given the strategy of his opponent, he can always choose a strategy in each stage so that he receives at least his minmax payoff. The possible equilibrium average payoffs of the infinitely repeated game are therefore bounded below by the pair (m_1, m_2), which in this game equals $(-2,-2)$:
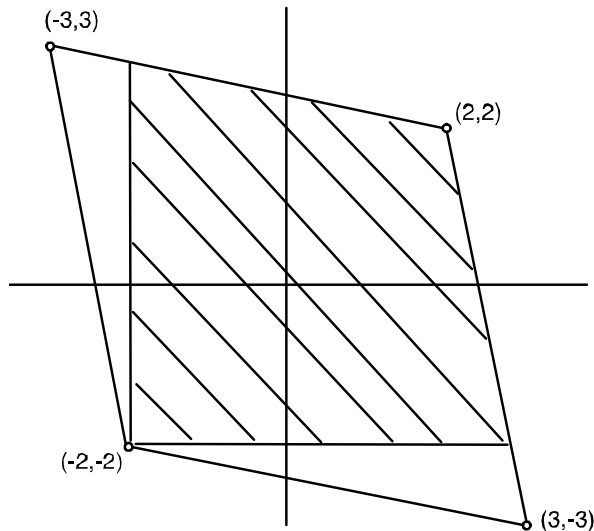
The Folk Theorem that we are discussing here states that **any** point (π_1^*, π_2^*) in this shaded region is the average payoff in some Nash equilibrium of the supergame.

The idea will be clear from considering a point (π_1^*, π_2^*) in which both entries are rational numbers. As rational numbers, the vector (π_1^*, π_2^*) can be written as

$$(\pi_1^*, \pi_2^*) = \rho_1(2,2) + \rho_2(-3,3) + \rho_3(3,-3) + \rho_4(-2,-2)$$

where each ρ_j is a nonnegative rational number such that

$$\rho_1 + \rho_2 + \rho_3 + \rho_4 = 1.$$

Let $n \in \mathbb{N}$ denote a common denominator for these four rational numbers. We consider a "cycle" $\mathcal{C}$ consisting of n stages in which each of the four outcomes $(2,2), (-3,3), (3,-3), (-2,-2)$ is played with the frequency determined by the numerator of ρ_i once the common demominator n has been chosen. The order in which the outcomes are played can be chosen arbitrarily, but the order is fixed and known to the players as part of their strategies. We consider the following trigger strategy for each player i:

Start the game by selecting the strategy specified by the first outcome in the cycle $\mathcal{C}$. Follow the cycle again and again unless at some stage the outcome specified by the cycle $\mathcal{C}$ is not played. In this case, switch to $\widetilde{s}_i$ ever after.

It is easy to see that this is a Nash equilibrium when players evaluate their sequence of payoffs in the infinitely repeated game using the limiting average method:

- Consider the prisoner's dilemma game above. By following the cycle, player i receives each of the four outcomes with frequencies as specified by the cycle, resulting in an average payoff of π_i^*. By deviating from the cycle, he receives at most m_i in every stage of some infinite tail (depending on when he deviates from the cycle), resulting in an average payoff of m_i. Because by assumption $m_i \leq \pi_i^*$ for each player i, the trigger strategies form a Nash equilibrium.

- Read over the last bullet point. It reflects the prisoner's dilemma only insofar as it mentions four outcomes of the game. A different convex hull would be drawn for each different stage game, and

(depending on the size of the game) a cycle might involve many more than four different outcomes. The principle, however, remains the same: because by assumption $m_i \leq \pi_i^*$ for each player i, the trigger strategies form a Nash equilibrium.

- In general, the trigger strategies that we have defined will not define a subgame perfect Nash equilibrium because the play of $(\widetilde{s}_1, \widetilde{s}_2)$ need not form a Nash equilibrium off the equilibrium path. In the case of the prisoner's dilemma, however, $(\widetilde{s}_1, \widetilde{s}_2) = (nc, nc)$, which is a Nash equilibrium of the stage game. The play of $(\widetilde{s}_1, \widetilde{s}_2) = (nc, nc)$ is a Nash equilibrium of the supergame with limiting average payoffs, and so the trigger strategies form an SPNE in this case.

- What if the objective (π_1^*, π_2^*) consists of a pair of irrational numbers? Consider a sequence of rational pairs
$$(\pi_1^t, \pi_2^t)_{t \in \mathbb{N}} \to (\pi_1^*, \pi_2^*).$$
For each t, one constructs a cycle $\mathcal{C}^t$ that implements (π_1^t, π_2^t) as the average payoffs over the play of $\mathcal{C}^t$. The trigger strategies then specify that the players move successively through $\mathcal{C}^1, \mathcal{C}^2, ...$ as the game is played, again with the threat of $(\widetilde{s}_1, \widetilde{s}_2)$ if any one ever fails to play according to the specified cycles. We will not pursue this more thoroughly because it seems to be mainly a point of mathematical interest.

- Returning to discounted payoffs, can we implement (π_1^*, π_2^*)? This is a little tricky because the discounting must be taken into account in selecting the cycle (i.e., the discounted payoffs over the cycle are (π_1^*, π_2^*)). But the result extends: if $(\pi_1^*, \pi_2^*) \geq (m_1, m_2)$, and if the discount factors of the traders are sufficiently large, then there exists a Nash equilibrium of the supergame whose discounted payoffs are (π_1^*, π_2^*).

- There are many versions of the folk theorem. They differ mainly by the solution concept used (i.e., what properties one wants the equilibrium to have that implements the particular payoffs). If you are interested in learning more about this topic, consult the text "Repeated Games and Reputations: Long-Run Relationships" by George Mailath and Larry Samuelson.