September 15

0.0.9 Finite Repetition of a Game

stage game: the game that is played repeatedly

$\mathbf{n}$ **stage game**: repetition of the game n times, with payoffs accumulating. For now, we won't address discounting.

Theorem 1 *If a game has a unique Nash equilibrium, then its finite repetition has a unique SPNE*

Proof. The proof is by induction on the number n of repetitions. Consider the n stage game.

The case of $n = 1$ is trivial.

In the nth repetition, consider each $n - 1$ stage game that follows as a subgame determined by the outcome of the first play of the game. Subgame perfection requires a subgame perfect NE in each subgame, and the induction hypothesis therefore uniquely characterizes the play of the game in stages 2 through n (regardless of the outcome of stage 1). The play of the game in stage 1 does not alter the play of the game in all subsequent stages. Consequently, the assumption of Nash equilibrium for the n-stage game requires the play of the unique Nash equilibrium in the first stage. ■

Notice that the result concerns not only the equilibrium path, but also the specification of the strategies off the equilibrium path.

Note also the independence of history that is described by the theorem. Long-term relationships ought to be different, but modeling an enduring relationship as a finite repetition of a game may fail produce interesting or different results. This is one of the reasons that so much of dynamic game theory focuses on infinitely repeated games.

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.,

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

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

Digression: The formula for the sum of a geometric series. Let $S = \sum_{t=0}^{\infty} \delta^t$ for $\delta \in [0,1)$. We have

$$\begin{aligned}\delta S &= \delta \cdot \sum_{t=0}^{\infty} \delta^t \\ &= \sum_{t=1}^{\infty} \delta^t \\ &= S - 1.\end{aligned}$$

Therefore,

$$(\delta - 1)\, S = -1 \Leftrightarrow S = \frac{1}{1-\delta}.$$

An alternative approach to discounting 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 sometimes easy to calculate the limiting payoff if the sequence of payoffs

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

eventually reaches some constant payoff,

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

regardless of the payoffs over stages 0 through $t^{'} - 1$.

Example 25 (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:

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

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 Nash equilibrium? 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:*

$$\begin{aligned}
(1-\delta_i)\sum_{t=0}^{\infty}\delta_i^t\cdot 2 &\geq (1-\delta_i)\left[\sum_{t=0}^{k-1}\delta_i^t\cdot 2+\delta_i^k\cdot 3+\sum_{t=k+1}^{\infty}\delta_i^t\cdot(-2)\right]\Leftrightarrow\\
\sum_{t=k}^{\infty}\delta_i^t\cdot 2 &\geq \delta_i^k\cdot 3+\sum_{t=k+1}^{\infty}\delta_i^t\cdot(-2)\Leftrightarrow\\
\sum_{t=k}^{\infty}\delta_i^t\cdot 2 &\geq \delta_i^k\cdot 3+\sum_{t=k+1}^{\infty}\delta_i^t\cdot(-2)\ \textit{cancel the first } k \textit{ terms}\\
&\Leftrightarrow \sum_{t=0}^{\infty}\delta_i^t\cdot 2\geq 3+\sum_{t=1}^{\infty}\delta_i^t\cdot(-2)\ \textit{cancel } \delta_i^k\\
\frac{2}{1-\delta_i} &\geq 3+\frac{\delta_i}{1-\delta_i}(-2)\\
2 &\geq 3-3\delta_i-2\delta_i\\
5\delta_i &\geq 1\\
\delta_i &\geq \frac{1}{5}
\end{aligned}$$

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 : *play nc in every stage*

   2 : *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.*

0.0.9.1 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

$$d_i > \pi_i^* > \pi_i. \tag{5}$$

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:

$$\begin{aligned}
(1-\delta_i)\sum_{t=0}^{\infty}\delta_i^t\cdot\pi_i^* &\geq (1-\delta_i)\left[\sum_{t=0}^{k-1}\delta_i^t\cdot\pi_i^* + \delta_i^k\cdot d_i + \sum_{t=k+1}^{\infty}\delta_i^t\cdot\pi_i\right] \Leftrightarrow \\
\sum_{t=k}^{\infty}\delta_i^t\cdot\pi_i^* &\geq \delta_i^k\cdot d_i + \sum_{t=k+1}^{\infty}\delta_i^t\cdot\pi_i \Leftrightarrow \\
\sum_{t=k}^{\infty}\delta_i^t\cdot\pi_i^* &\geq \delta_i^k\cdot d_i + \sum_{t=k+1}^{\infty}\delta_i^t\cdot\pi_i \text{ \textit{cancel the first k terms}} \\
&\Leftrightarrow \sum_{t=0}^{\infty}\delta_i^t\cdot\pi_i^* \geq d_i + \sum_{t=1}^{\infty}\delta_i^t\cdot\pi_i \text{ \textit{cancel} } \delta_i^k \\
\frac{\pi_i^*}{1-\delta_i} &\geq d_i + \frac{\delta_i}{1-\delta_i}\cdot\pi_i \\
\pi_i^* &\geq d_i(1-\delta_i) + \delta_i\pi_i
\end{aligned}$$

$$\begin{aligned}
\pi_i^* - d_i &\geq \delta_i (\pi_i - d_i) \\
d_i - \pi_i^* &\leq \delta_i (d_i - \pi_i) \\
\delta_i &\geq \frac{d_i - \pi_i^*}{d_i - \pi_i} \text{ applying (5)}
\end{aligned}$$

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 26 *Consider 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*

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.