1 Repeated Games

In this section we will discuss repeated games. For a repeated game we generally take a game that is played simultaneously, such as the Prisoner's Dilemma or the Boxing-Opera game, and have the players play that game repeatedly. The basic game that is played is called the *stage game* or the one-shot game. So if the Prisoner's Dilemma is repeated two times we would call the Prisoner's Dilemma the stage game.

One goal of repeated games is to model repeated actions between the players. A second goal is to determine if different equilibria arise when the game is repeated. For instance, one might imagine that two individuals playing the Prisoner's Dilemma repeatedly might choose Not Confess at times as they know that they will play this game again. We will focus on two types of repeated games – finitely repeated games and infinitely repeated games. There are differences between the two.

2 Finitely Repeated Games

A finitely repeated game is one that is repeated a finite number of times. It does not matter how many times, as long as the players know the exact number. It could be a game that is played twice or it could be a game that is played 1,000,000 times. As long as the exact number is known it is a finite game. We will consider the Prisoner's Dilemma that is played twice. We will use the Prisoner's Dilemma game with the firms making quantity choices as this is a game with positive payoffs. The stage game is:

		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	\$11, \$11	\$3, \$16
	$Q = 20$	\$16, \$3	\$5, \$5

Recall that both players in this game have a strictly dominant strategy to choose $Q = 20$. However, they would both be better off if they would both choose $Q = 10$. This notion of them both being better off if they could choose $Q = 10$ is amplified if the game is played twice. Now if they choose $Q = 10$ both times they receive \$22 whereas by choosing $Q = 20$ both times they only receive \$10. However, just because they are better off (jointly) choosing $Q = 20$ does not make this a Nash equilibrium. We need to check to make sure that both are playing a best response at all subgames.

Now, why might cooperating work when games are repeated? There is always a threat of punishment when the games are repeated – if Firm A does not cooperate in period 1, then Firm B can threaten to punish in period 2 (or later periods if they exist in the game). The question is how credible are these threats.

2.1 Game Tree

Figure 1 shows the twice repeated Prisoner's Dilemma. Note that in this game there are 5 subgames. One subgame is the entire game. The other 4 subgames are the repeated games after the first subgame. So there is one subgame after the firms choose $(10, 10)$ in the first round; a second subgame after the firms choose $(10, 20)$ in the first round; a third subgame after the firms choose $(20, 10)$ in the first round; and a fourth subgame after the firms choose $(20, 20)$ in the first round. The first number in parentheses refers to firm A's choice. The final payoffs are found by adding the payoffs from each round of play. For instance, if firms choose $(10, 10)$ in the first round then they each receive \$11 in the first round. If they choose $(10, 20)$ in the first round then Firm A receives \$3 and Firm B receives \$16 in the first round. If the choice is $(20, 10)$ then Firm A receives \$16 and Firm B receives \$3 in the first round. Finally, if the firms choose $(20, 20)$ in the first round then each receives \$5. Note that the firms are not conspiring to make their choices nor are they communicating in any way. This is just a "what if the firms chose" exercise. After the first round is played there are still four possible outcomes that could occur. The table below shows how the final payoffs are calculated:

Firm	A	B	Tot Pay A, B	A	B	Tot Pay A, B	A	B	Tot Pay A, B	A	B	Tot Pay A, B
1st round action	10	10		10	20		20	10		20	20	
1st round payoff	11	11		3	16		16	3		5	5	
$(10,10)$ 2nd, pay	11	11	$22, 22$	11	11	$14, 27$	11	11	$27, 14$	11	11	$16, 16$
$(10,20)$ 2nd, pay	3	16	$14, 27$	3	16	$6, 32$	3	16	$19, 19$	3	16	$8, 21$
$(20,10)$ 2nd, pay	16	3	$27, 14$	16	3	$19, 19$	16	3	$32, 6$	16	3	$21, 8$
$(20,20)$ 2nd, pay	5	5	$16, 16$	5	5	$8, 21$	5	5	$21, 8$	5	5	$10, 10$

The columns for "Tot Pay A,B" simply show the total payoff for Firms A and B given their first round actions and their second round actions.

Now, how do we solve this game? We simply start from the smallest subgame (or one of the smallest) and work back to the beginning. Suppose that the firms both choose $Q = 10$ in the first round. The strategic (or normal or matrix) form of the subgame that follows that set of choices is:

If $(10, 10)$ is the outcome of round 1

		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	$22, 22$	$14, 27$
	$Q = 20$	$27, 14$	$16, 16$

Note that the only Nash equilibrium to this subgame is Firm A choose 20 and Firm B choose 20. Now consider Firm A choosing 10 in round 1 and Firm B choosing 20 in round 1. The strategic form of this subgame is:

If $(10, 20)$ is the outcome of round 1

		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	$14, 27$	$6, 32$
	$Q = 20$	$19, 19$	$8, 21$

Note that the only Nash equilibrium to this subgame is Firm A choose 20 and Firm B choose 20. Now consider Firm A choosing 20 in round 1 and Firm B choosing 10 in round 1. The strategic form of this subgame is:

If $(20, 10)$ is the outcome of round 1

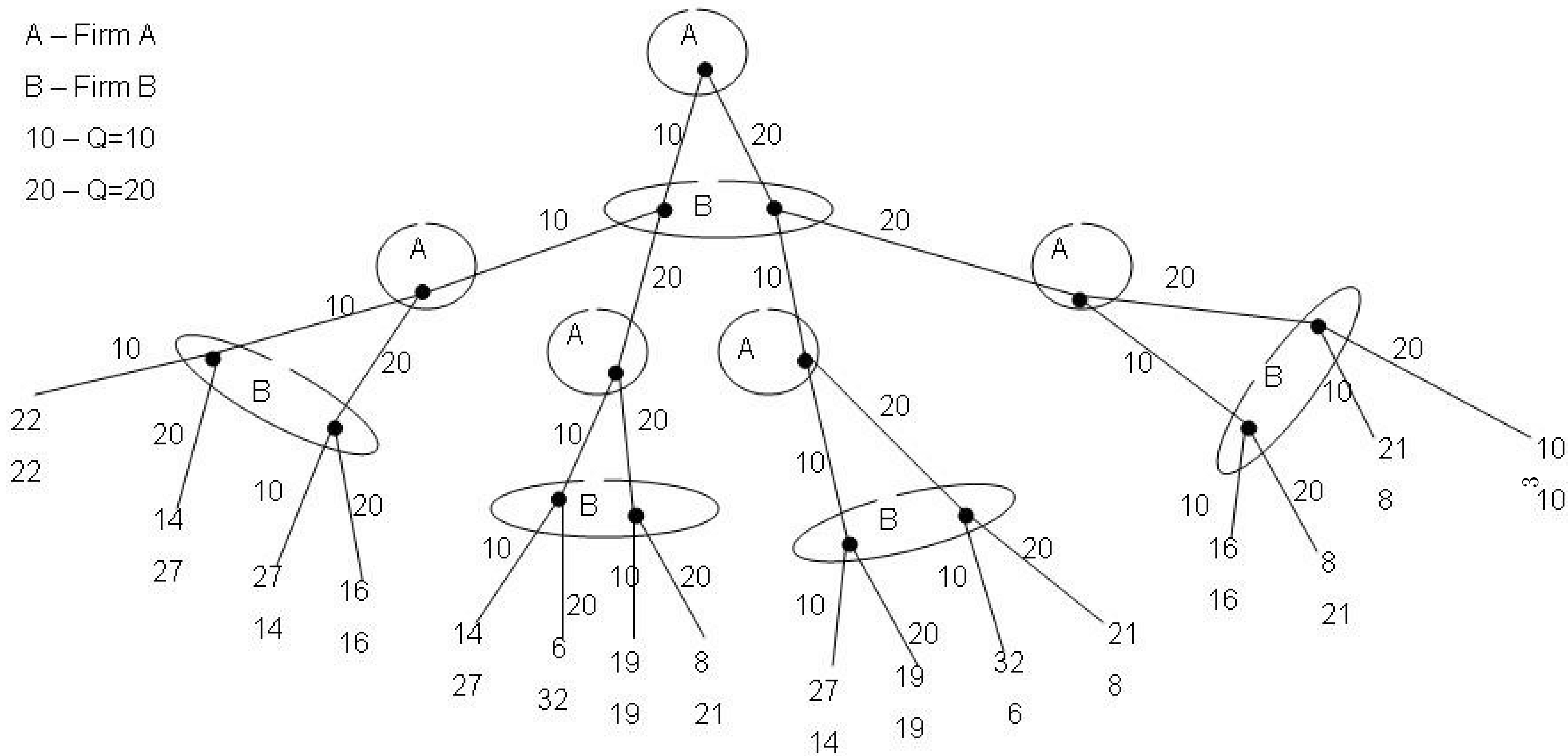
		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	$27, 14$	$19, 19$
	$Q = 20$	$32, 6$	$21, 8$

Note that the only Nash equilibrium to this subgame is Firm A choose 20 and Firm B choose 20. Now consider Firm A choosing 20 in round 1 and Firm B choosing 20 in round 1. The strategic form of this subgame is:

If $(20, 20)$ is the outcome of round 1

		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	$16, 16$	$8, 21$
	$Q = 20$	$21, 8$	$10, 10$

Once again, note that the only Nash equilibrium to this subgame is Firm A choose 20 and Firm B choose 20. In fact, choosing 20 is a strictly dominant strategy for all of these subgames. Given that we now know what the firms will do at each subgame we can reduce the original game (the entire game where the Prisoner's Dilemma is played twice) to a much simpler game. If $(10, 10)$ is played in round 1 then we know that the firms will both choose $Q = 20$ in round 2 and that the payoffs will be 16 for Firm A and 16 for Firm B. If $(10, 20)$ is played in round 1 then we know that the firms will both choose $Q = 20$ in round 2 and that the payoffs will be 8 for Firm A and 21 for Firm B. If $(20, 10)$ is played in round 1 then we know that the firms will both choose $Q = 20$ in round 2 and that the payoffs will be 21 for Firm A and 8 for Firm B. If $(20, 20)$ is played in round 1 then we know that the firms will both choose $Q = 20$ in round 2 and that the payoffs will be 10 for Firm A and 10 for Firm B. Thus, the new game tree (eliminating all the branches that will not be played) is in Figure 2.The strategic form of this game is:


A – Firm A
B – Firm B
10 – Q=10
20 – Q=20

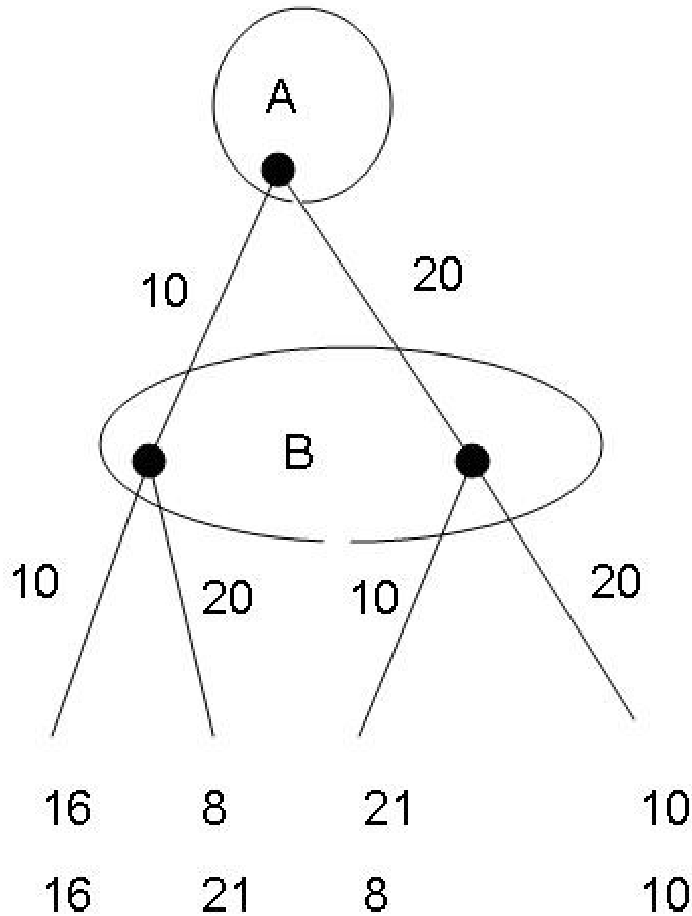


Figure 2: Repeated Prisoner's Dilemma with the 2nd round eliminated

Reduced version of twice played Prisoner's Dilemma

		Firm B	
		$Q = 10$	$Q = 20$
Firm A	$Q = 10$	$16, 16$	$8, 21$
	$Q = 20$	$21, 8$	$10, 10$

Note that in this game there is only one Nash equilibrium, both firms choose $Q = 20$. Given what we already know about what should occur in the second round, we now see that both firms choosing $Q = 20$ whenever they have a chance to make a decision is the SPNE of this game. Formally the strategy would look something like:

Firm A choose 20 in the first period and Firm B choose 20 in the first period. In the second period:

	Firm A		Firm B
	choose		choose
if $(10, 10)$ in 1st period	$Q = 20$		$Q = 20$
if $(10, 20)$ in 1st period	$Q = 20$		$Q = 20$
if $(20, 10)$ in 1st period	$Q = 20$		$Q = 20$
if $(20, 20)$ in 1st period	$Q = 20$		$Q = 20$

Note that this specifies an action at each information set for each player. But so does "Both players always choose 20 whenever they have the opportunity to make a decision". The key is the word "always" – just stating that Firm A should choose 20 and Firm B should choose 20 is NOT a complete strategy because it is unclear what that means.

The outcome to this game is that both firms end up receiving \$10, so they are unable to cooperate in either period even when the game is played twice. The goal was to see if repeating the game could lead to more cooperation (choosing $Q = 10$) but it did not, at least not for 2 periods. Think about the Prisoner's Dilemma played 3 times. Would this allow the firms to cooperate in any period? How about 4, 5, or 10 times? Would this allow the firms to cooperate in any period? Always start at the end of the game and work towards the beginning. Think about a 10 period game. In the 10^{th} period the only Nash equilibrium to th0se subgames is for both firms to choose $Q = 20$. The reason is that there are no future periods in which either firm can punish the other firm if it chooses $Q = 20$. Now consider the 9^{th} period. We know that the firms will both choose $Q = 20$ in period 10, so if one firm does not cooperate in period 9 there is no means of punishing that firm (choosing $Q = 10$ when the other player chooses $Q = 20$ is NOT a punishment). Now think about the 8^{th} period – again, no means of punishing the other firm. This continues all the way back to the 1^{st} period and, in equilibrium, the firms end up choosing $Q = 20$ at every information set they have in the Prisoner's Dilemma type games as long as the game is *FINITELY* repeated. In fact, when all players have a strictly dominant strategy in the stage game, which $Q = 20$ is, there is a theorem that states that the only SPNE to the finitely repeated game is the one where all players choose the strictly dominant strategy whenever they get to make a decision.

Theorem 1 *In finitely repeated games, if all players have a strictly dominant strategy in the stage game then the unique SPNE to the game is for all players to choose their strictly dominant strategy whenever they have a chance to make a decision.*

Note that while this may be an unsatisfying answer to whether or not the firms are able to cooperate when the game is repeated a finite amount of times it is the only SPNE. However, there are other types of games where there is not a strictly dominant strategy for all players in the stage game. One type is a game like Boxing-Opera (a coordination game in general). But in this type of game nothing truly interesting happens – while there are multiple SPNE, the players always choose to meet up in both stages of the game; the only question is whether they decide on Boxing in round 1 and Opera in round 2, or always Boxing, or always Opera, or Opera in round 1 and Boxing in round 2. With 3x3 (and bigger) stage games it is possible that a SPNE to the game involves an outcome in the stage game that is NOT a NE to the stage game. For example, consider this game:

		Player 2				
		L_2	M_2	R_2	P_2	Q_2
	L_1	$1,1$	$5,0$	$0,0$	$0,0$	$0,0$
	M_1	$0,5$	$4,4$	$0,0$	$0,0$	$0,0$
Player 1	R_1	$0,0$	$0,0$	$3,3$	$0,0$	$0,0$
	P_1	$0,0$	$0,0$	$0,0$	$4,\frac{1}{2}$	$0,0$
	Q_1	$0,0$	$0,0$	$0,0$	$0,0$	$\frac{1}{2},4$

Note that the outcome $(4,4)$, which occurs when Player 1 uses M_1 and Player 2 uses M_2, is NOT a Nash equilibrium to this stage game. However, that outcome cell is payoff dominant to all of the PSNE in the game (L_1, L_2); (R_1, R_2); (P_1, P_2); and (Q_1, Q_2). It is possible, if this game is played twice, for the (M_1, M_2) set of strategies to be played as part of a SPNE in the first stage of this game. While we will not go through the details, this example is from the Gibbons book mentioned in the syllabus, pages 87-88.

3 Infinitely Repeated Games

One may ask why we bother to study infinitely repeated games since most players of games outside of Duncan MacLeod tend to expire at some point in time. There are a few good reasons to study infinitely repeated games. One is that not all players need to expire – consider a corporation. Corporations can be infinitely lived and they play many economic games. A second reason is that although we will all eventually expire the endpoint of the game is (hopefully) uncertain. There are results that show that games that are repeated finitely with an uncertain endpoint are consistent with games that are infinitely repeated. The experimental evidence on play in finitely repeated games suggests a third reason. The subgame perfect Nash equilibrium is a poor predictor of behavior in some of these repeated games. Studying infinitely repeated games allows for a different set of SPNE to be chosen. There is one major drawback to infinitely repeated games, and that is that multiple equilibria (even multiple SPNE) are bound to exist. Some view this as a problem, but it just shifts the focus from trying to show that an equilibrium exists to trying to show why one of the equilibria should be selected over another.

3.1 Evaluating strategies in infinite games

We begin with the Prisoner's Dilemma game we have been using:

		Player 2	
		$Q = 10$	$Q = 20$
Player 1	$Q = 10$	$11, 11$	$3,16$
	$Q = 20$	$16,3$	$5,5$

It will be easier to relabel the strategies. Let $Q = 10$ be Cooperate and $Q = 20$ be Defect. Thus, the game is now:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$11, 11$	$3,16$
	Defect	$16,3$	$5,5$

Both players would be better off if they could both choose Cooperate instead of both choosing Defect, but Defect is a dominant strategy. In order to evaluate strategies in infinite games it will be necessary to add a particular parameter to the discussion. The parameter added will be the player's discount rate, δ. It is assumed that $\delta \in [0,1)$, and that players have exponential discounting. All that exponential discounting means is that a payoff one time period from today is discounted at δ and a payoff two time periods from today is discounted at δ^2, etc. Thus, a player's payoff stream from the infinite game would look like:

$$\delta^0\Pi_0 + \delta^1\Pi_1 + \delta^2\Pi_2 + \delta^3\Pi_3 + ...$$

where Π_k denotes the player's payoff in each period k. The $\delta \in [0,1)$ assumption will be justified shortly.[1] It is typically assumed that players (and people in general) prefer \$1 today to \$1 tomorrow, and \$1 tomorrow

[1] The exponential discounting assumption is used because it allows for time consistent preferences. Hyperbolic discounting is another type of discounting that has been suggested as consistent with choices made by individuals in experiments, although hyperbolic discounting does not necessarily lead to time consistent preferences.

to \$1 two days from now. Thus, the sooner a player receives a payoff the less he discounts it. Why add this discount rate? Well, if we do not have a discount rate then the players' payoffs from following ANY strategy (assuming that there are no negative payoffs that the player could incur) of an infinite game would be infinite. That's not very interesting. This is also why we assume that $\delta < 1$ rather than $\delta \leq 1$. If $\delta = 1$, then a player weights all payoffs equally regardless of the time period, and this leads to an infinite payoff. If $\delta = 0$, then the player will only care about the current period. As δ moves closer to 1, the player places more weight on future periods. It is possible to motivate this discount rate from a present value context, which would make $\delta = \frac{1}{1+r}$, where r is "the interest rate". Thus, if $r = 0.05$, then $\delta \approx 0.95$. All this says is that getting \$1 one period from today is like getting 95 cents today, and getting \$1 two periods from today is like getting 90.7 cents today. While this interpretation of the discount rate is the most closely linked to economic behavior, we will not assume that the discount rate is directly related to the interest rate, but that it is simply a parameter that states how players value payoffs over time.

Now, suppose that players 1 and 2 use the following strategies:

Player 1 chooses Cooperate in the initial period (at time $t = 0$) and continues to choose Cooperate at every decision node unless he observes that player 2 has chosen Defect. If Player 1 ever observes Player 2 choosing Defect then Player 1 will choose Defect at every decision node after that defection. Player 2's strategy is the same. These strategies call for Cooperation at every decision node until a Defection is observed and then Defection at every decision node after Defection is observed. Note that this is a *potential* SPNE because it is a set of strategies that specifies an action at every decision node of the game. The question then becomes whether or not this is a SPNE of the game. Recall that a strategy profile is an SPNE if and only if it specifies a NE at every subgame. Although each subgame of this game has a distinct history of play, all subgames have an identical structure. Each subgame is an infinite Prisoner's Dilemma exactly like the game as a whole. To show that these strategies are SPNE, we must show that after any previous history of play the strategies specified for the remainder of the game are NE.

Consider the following two possibilities:

1. A subgame that contains a deviation from the Cooperate, Cooperate outcome somewhere prior to the play of the subgame

2. A subgame that does not contain a deviation from the Cooperate, Cooperate outcome

If a subgame contains a deviation then the players will both choose Defect, Defect for the remainder of the game. Since this is the NE to the stage game of the Prisoner's Dilemma, it induces a NE at every subgame. Thus, the "Defect if defection has been observed" portion of the suggested strategy induces NE at every subgame.

Now, for the more difficult part. Suppose that the players are at a subgame where no previous defection has occurred. Consider the potential of deviation from the proposed strategy in period $\tau \geq t$, where t is the current period. If player 2 chooses Defect in period τ he will earn $\delta^{\tau}\Pi^{Deviate} + \delta^{\tau}\sum_{t=1}^{\infty}\delta^{t}\Pi^{D}$ for the remainder of the game, where $\Pi^{Deviate}$ is player 2's payoff from deviating and Π^{D} is his payoff each period from the (Defect, Defect) outcome. If player 2 chooses to follow the proposed strategy, then he will earn $\delta^{\tau}\sum_{t=0}^{\infty}\delta^{t}\Pi^{C}$, where Π^{C} is his payoff from the (Cooperate, Cooperate) outcome. The question then becomes under what conditions will the payoff from following the proposed strategy be greater than that from the payoff of deviating. To find the condition simply set up the inequality:

$$\delta^{\tau}\sum_{t=0}^{\infty}\delta^{t}\Pi^{C} \geq \delta^{\tau}\Pi^{Deviate} + \delta^{\tau}\sum_{t=1}^{\infty}\delta^{t}\Pi^{D}$$

We can cancel out the δ^{τ} terms to obtain:[2]

$$\sum_{t=0}^{\infty}\delta^{t}\Pi^{C} \geq \Pi^{Deviate} + \sum_{t=1}^{\infty}\delta^{t}\Pi^{D}$$

Now, using results on series from Calculus,[3] we have:

$$\frac{1}{1-\delta}\Pi^{C} \geq \Pi^{Deviate} + \frac{\delta}{1-\delta}\Pi^{D}$$

[2] This canceling out of the δ^{τ} terms typically leads to the assumption that if deviation is going to occur in an infinitely repeated game it will occur in the first time period. I proceed under this assumption in later examples.

[3] See the following subsection for more detail on how we moved from one step to the other.

Now, we can substitute in for $\Pi^{Deviate}, \Pi^D$, and Π^C from our game to find:

$$11\frac{1}{1-\delta} \geq 16 + 5\frac{\delta}{1-\delta}$$

Or:

$$\begin{aligned} 11 &\geq 16 - 16\delta + 5\delta \\ 0 &\geq 5 - 11\delta \\ 11\delta &\geq 5 \\ \delta &\geq \frac{5}{11} \end{aligned}$$

Thus, choosing to deviate from the proposed strategy only provides a higher payoff if $\delta \leq \frac{5}{11}$, so that continuing to cooperate is a best response if $\delta \geq \frac{5}{11}$. Thus, in this game, as long as both players are patient enough (with $\delta \geq \frac{5}{11}$) they will end up choosing Cooperate at every decision node. This is a much better outcome for the players than choosing Defect at every decision node.

The discount rate will be a key factor in determining whether or not a proposed equilibrium is a SPNE. In fact, when looking at infinitely repeated games, it is best to have a particular strategy in mind and then check to see what the necessary conditions are for it to be a SPNE, given the multiplicity of equilibria.

3.1.1 Digression on series

In the section above we moved from:

$$\begin{aligned} \textstyle\sum_{t=0}^{\infty} \delta^t \Pi^C &\geq \Pi^{Deviate} + \textstyle\sum_{t=1}^{\infty} \delta^t \Pi^D \\ &\quad \text{to} \\ \frac{1}{1-\delta}\Pi^C &\geq \Pi^{Deviate} + \frac{\delta}{1-\delta}\Pi^D \end{aligned}$$

How was this done? First, think about what $\sum_{t=0}^{\infty} \delta^t \Pi^C$ is:

$$\begin{aligned} \textstyle\sum_{t=0}^{\infty} \delta^t \Pi^C &= \Pi^C + \delta\Pi^C + \delta^2\Pi^C + ... + \delta^{\infty}\Pi^C \\ \textstyle\sum_{t=0}^{\infty} \delta^t \Pi^C &= \Pi^C\left(1 + \delta + \delta^2 + ... + \delta^{\infty}\right) \end{aligned}$$

When $\delta \in [0, 1)$ this infinite series becomes $\frac{1}{1-\delta}$. We will not prove this result but we will use the result. That explains the right hand side of the inequality. Now for the left hand side, although this is just a variation on the right hand side. We have $\Pi^{Deviate} + \sum_{t=1}^{\infty} \delta^t \Pi^D$:

$$\Pi^{Deviate} + \textstyle\sum_{t=1}^{\infty} \delta^t \Pi^D = \Pi^{Deviate} + \delta\Pi^D + \delta^2\Pi^D + \delta^3\Pi^D + ... + \delta^{\infty}\Pi^D$$

The $\Pi^{Deviate}$ terms are easy – they are just the same. Notice the difference in the summation of the Π^D and the Π^C terms. The summation of the Π^C terms starts at $t = 0$ while the summation of the Π^D terms begins at $t = 1$. This is because the summation of the Π^D terms starts one period later. If we focus solely on $\sum_{t=1}^{\infty} \delta^t \Pi^D$ we have:

$$\begin{aligned} \textstyle\sum_{t=1}^{\infty} \delta^t \Pi^D &= \delta\Pi^D + \delta^2\Pi^D + \delta^3\Pi^D + ... + \delta^{\infty}\Pi^D \\ \delta\Pi^D + \delta^2\Pi^D + \delta^3\Pi^D + ... + \delta^{\infty}\Pi^D &= \delta\Pi^D\left(1 + \delta + \delta^2 + ... + \delta^{\infty}\right) \end{aligned}$$

But the term $1 + \delta + \delta^2 + ... + \delta^{\infty}$ is just the infinite sum, so we now have:

$$\delta\Pi^D\left(1 + \delta + \delta^2 + ... + \delta^{\infty}\right) = \frac{\delta\Pi^D}{1-\delta}$$

because:

$$1 + \delta + \delta^2 + ... + \delta^{\infty} = \frac{1}{1-\delta}.$$

This is how we obtain those results above.

3.2 Multiple SPNE

Are there other SPNE to the game? Yes, there are many SPNE to this game. In games like the Prisoners' Dilemma, which have a strictly dominant strategy in the stage game, both players choosing to use their strictly dominant stage game strategies whenever they get to make a decision is *always* an SPNE to the game. But there are many others. Consider a modified version of the game:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	11, 11	3, 32
	Defect	32, 3	5, 5

The only change in this game is that the payoff of 16 that the player received from Defecting when the other player Cooperates has been changed to 32. We can show that both players using a strategy of cooperating until a defection occurs (the same proposed strategy from before) is a SPNE if:

$$11\frac{1}{1-\delta} \geq 32 + 5\frac{\delta}{1-\delta}$$

or $\delta \geq \frac{7}{9}$. Thus, if both players are sufficiently patient then the proposed strategy is still a SPNE. Note that the discount rate increased in this example since the payoff to deviating increased. But, is there a strategy that yields higher payoffs than both players receiving 11 each period? What if the following strategies were used by players 1 and 2:

If no deviation has occurred, Player 1 chooses Defect in all even time periods and chooses Cooperate in all odd time periods. If a deviation occurs Player 1 always chooses Defect. Player 1 chooses Defect at time $t = 0$ and Cooperate at time $t = 1$.

If no deviation has occurred, Player 2 chooses Cooperate in all even time periods and chooses Defect in all odd time periods. If a deviation occurs Player 2 always chooses Defect. Player 2 chooses Cooperate at time $t = 0$ and Defect at time $t = 1$.

A deviation (from player 1's perspective) occurs when Player 2 chooses Defect in an even time period. A deviation (from player 2's perspective) occurs when Player 1 chooses Defect in an odd time period. Note that we start the game at time $t = 0$, so that Player 1 receives 32 first.

Look at what this strategy would do. It would cause the outcome of the game to alternate between the $(Defect, Cooperate)$ and $(Cooperate, Defect)$ outcomes, giving the players alternating periods of payoffs of 32 and 3, as opposed to 11 each period using the "cooperate until defect is observed, then always defect" strategy. On average (and ignoring discounting for a moment), each player would receive 17.5 per period under this new strategy and only 11 per period under the old. Is the new strategy a SPNE? We should check for both players now that they are receiving different amounts of payoffs in different periods.

For Player 1:

$$\begin{aligned}
\Pi^{Deviate} &= 32 + \textstyle\sum_{t=1}^{\infty} \delta^t 5 \\
\Pi^{C} &= \textstyle\sum_{t=0}^{\infty} \delta^{2t} 32 + \sum_{t=0}^{\infty} \delta^{2t+1} 3
\end{aligned}$$

If $\Pi^C \geq \Pi^{Deviate}$ then Player 1 will choose NOT to deviate:

$$\begin{aligned}
32\frac{1}{1-\delta^2} + 3\frac{\delta}{1-\delta^2} &\geq 32 + 5\frac{\delta}{1-\delta} \\
32 + 3\delta &\geq 32 - 32\delta^2 + 5\delta\,(1+\delta) \\
32 + 3\delta &\geq 32 - 32\delta^2 + 5\delta + 5\delta^2 \\
3\delta &\geq -32\delta^2 + 5\delta + 5\delta^2 \\
0 &\geq 2\delta - 27\delta^2 \\
27\delta^2 - 2\delta &\geq 0 \\
27\delta - 2 &\geq 0 \\
\delta &\geq \frac{2}{27}
\end{aligned}$$

For Player 2 we will get the exact same number:

$$\begin{aligned} \Pi^{Deviate} &= \textstyle\sum_{t=0}^{\infty} \delta^i 5 \\ \Pi^{C} &= \textstyle\sum_{t=0}^{\infty} \delta^{2i} 3 + \sum_{t=0}^{\infty} \delta^{2i+1} 32 \end{aligned}$$

If $\Pi^C \geq \Pi^{Deviate}$ then Player 2 will choose NOT to deviate:

$$\begin{aligned} 3\frac{1}{1-\delta^2} + 32\frac{\delta}{1-\delta^2} &\geq 5\frac{1}{1-\delta} \\ 3 + 32\delta &\geq 5 + 5\delta \\ 27\delta &\geq 2 \\ \delta &\geq \frac{2}{27} \end{aligned}$$

Thus, both players need to have a discount rate greater than or equal to $\frac{2}{27}$ to support this strategy. Note that this discount rate is much lower than the one needed to support the "cooperate until defect is observed, then always defect" strategy. However, it also illustrates the "embarrassment of riches" of infinitely repeated games because for any $\delta \geq \frac{7}{9}$ either of these strategies could be played. And those are NOT the only two strategies.

3.2.1 Some results

There are a number of formal results for SPNE that one can show concerning infinite games. Most of these results hinge upon a discount rate δ being sufficiently close to 1. In the Prisoner's Dilemma type games we have considered the punishment for deviating from the "cooperation" strategy is for the other player to play the stage game (or single shot) Nash equilibrium for the remainder of the game (choose Defect forever). Since there are only two actions a player can take at any decision node (Cooperate or Defect) the only method of punishment is to play Defect. Equilibria where the punishment takes the form of playing the stage game Nash equilibrium are known as Nash reversion since the game reverts back to the Nash equilibrium once a defection is observed.

Supporting average payoffs greater than stage game Nash As we showed in the second example (when the players alternated choosing the Cooperate and Defect strategies) it need not be the case that the players always "agree"[4] to choose the same strategy in each period. It is possible to show that ANY payoff stream that yields average (undiscounted) payoffs above the Nash equilibrium level can be supported by the threat of Nash reversion IF the discount rate is sufficiently close to 1. Again, consider a (modified) Cooperate, Defect game:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	25, 25	4, 80
	Defect	80, 4	8, 8

Suppose that you believe that Player 1 is a much tougher player than Player 2, whatever being "tougher" means. A potential SPNE strategy is as follows: Player 2 always chooses Cooperate unless a deviation is observed and then chooses Defect forever once deviation is observed. Exactly what a deviation is will be made clear momentarily, but consider Player 1 who plays Cooperate in the first period, then Defect for 4 periods, then Cooperate in the 6^{th} period, then Defect for 4 periods, etc., unless a defection is observed. A deviation by Player 2 is any play of Defect. A deviation by Player 1 is a choice of Defect in either the 1^{st}, 6^{th}, 11^{th}, etc. periods. If Player 1 were to choose Cooperate in periods 2, 3, 4, 5, 7, 8, 9, 10, etc. Player 2 would NOT view this as a deviation. Note that the payoff stream for Player 2 is 25, 4, 4, 4, 4, 25, 4, 4, 4, 4, etc. Every 5 periods Player 2 receives an undiscounted payoff of 41, with an average payoff of 8.2 per period. Since this average is greater than the payoff from playing the stage game Nash equilibrium (8),

[4] This is another slight problem with infinite games – there are so many SPNE that it is difficult to say how one particular one arose.

Player 2 will "agree" to play this equilibrium IF his discount rate is close enough to 1. Given that we have actual payoffs we can see that Player 2 will choose to always Cooperate if:

$$\sum_{t=0}^{\infty}\left(\delta^5\right)^t 25+\delta\sum_{t=0}^{\infty}\left(\delta^5\right)^t 4+\delta^2\sum_{t=0}^{\infty}\left(\delta^5\right)^t 4+\delta^3\sum_{t=0}^{\infty}\left(\delta^5\right)^t 4+\delta^4\sum_{t=0}^{\infty}\left(\delta^5\right)^t 4\geq 80+\delta\sum_{t=0}^{\infty}\delta^t 8$$

You should check that the payoff stream on the left-hand side is the actual payoff stream. Simplifying this expression gives:

$$\frac{25}{1-\delta^5}+\frac{4\delta}{1-\delta^5}+\frac{4\delta^2}{1-\delta^5}+\frac{4\delta^3}{1-\delta^5}+\frac{4\delta^4}{1-\delta^5}\geq 80+\frac{8\delta}{1-\delta}$$

This is not an easy equation to solve for δ, so we can evaluate it numerically. The goal (for the example) is not to find the actual discount rate but to show that for some discount rate close to 1 that this set of strategies constitutes a SPNE. Suppose that $\delta=0.99$. The left-hand side (Cooperate) is:

$\left[\frac{25}{1-\delta^5}+\frac{4\delta}{1-\delta^5}+\frac{4\delta^2}{1-\delta^5}+\frac{4\delta^3}{1-\delta^5}+\frac{4\delta^4}{1-\delta^5}\right]_{\delta=.99}=828.48$

The right-hand side (Defect) is:

$\left[80+\frac{8\delta}{1-\delta}\right]_{\delta=.99}=872.0$

So even for a $\delta=.99$ Player 2 would not play this SPNE. But 0.99 is not as close as we can get to 1. What if $\delta=0.999$? The left-hand side is:

$\left[\frac{25}{1-\delta^5}+\frac{4\delta}{1-\delta^5}+\frac{4\delta^2}{1-\delta^5}+\frac{4\delta^3}{1-\delta^5}+\frac{4\delta^4}{1-\delta^5}\right]_{\delta=.999}=8208.4$

While the right-hand side is:

$\left[80+\frac{8\delta}{1-\delta}\right]_{\delta=.999}=8072.0.$

So, for some discount rate between 0.99 and 0.999 this set of strategies becomes a potential solution. Another way to think about this is to consider the case where $\delta=1$. Now, since the game is played infinitely any set of strategies will lead to an infinite payoff, but it may be that one set of strategies gets to infinity "faster". Consider the first 355 periods of the game. Using the "defection strategy", Player 2 will have received 80 in the first period and 8 for the next 354 periods. This leads to an undiscounted payoff of 2912 for the 355 periods. Using the "cooperation strategy", Player 2 will have received an average payoff of 8.2 each period for the 355 periods for an undiscounted payoff of 2911. In the 356^{th} period Player 2 would receive 8 using the "defection strategy", bringing the total payoff up to 2920, and using the "cooperation strategy" will receive 25, bringing the total payoff up to 2936. Up until the 356^{th} period the total payoff from "defection" is less than the total payoff from "cooperation", but from the 356^{th} period onward the total payoff from "cooperation" is ALWAYS greater than or equal to that from "defection". From period 361 onward the total payoff from "cooperation" is ALWAYS strictly greater than that from "defection". In a sense, the "cooperation strategy" overtakes the "defection strategy" at some point in time and from that point in time onward is NEVER overtook by the "defection" strategy.

Supporting average payoffs less than stage game Nash It is also possible to support payoffs LESS than the stage game Nash equilibrium payoffs. However, this cannot happen in our Cooperate, Defect game because the minimum payoff a player can guarantee himself in that game is 8, which is the stage game Nash payoff (if someone plays Defect this guarantees that person will receive at least 8). It might be the case that punishment can be WORSE than the Nash equilibrium to the stage game (again, to be clear, this is not the case in the Cooperate, Defect game, but it could be in some other game). Thus it is possible to support average payoffs that are less than the stage game Nash equilibrium payoffs as a SPNE as long as the discount rate is close to 1. For example, if the Nash equilibrium payoff is 10, but the highest amount a player can guarantee himself is 6, then it is possible to find a SPNE where that player receives an average of 6.1 each period.

Carrot-and-stick approach So far all of our SPNE have used what is known as a "grim trigger" strategy. When using a grim trigger strategy, once a defection is observed play reverts to the Nash equilibrium (or worse) – forever. This is an extremely harsh punishment as it allows no room for error. An alternative is to use a carrot-and-stick approach. The punishment portion of the strategy specifies that the punisher will only punish for x periods rather than every ensuing period. That is the stick. The carrot is the cooperation payoff that the player receives x periods in the future once a defection is observed, provided the defecting player returns to cooperating. This approach is more forgiving than the grim trigger strategy, and in games where (1) mistakes may be made (2) actions may be misinterpreted or (3) there is some uncertainty that influences the players' payoffs in addition to the players' chosen actions this more forgiving approach may yield higher payoffs than the grim trigger strategy. A good example of this approach can be found in Green and Porter (1984),[5] *Econometrica*, Noncooperative Collusion Under Imperfect Price Information, 87-100. In that model there is a group of firms who wish to collude. The market price is influenced by the total quantity produced by each firm. In addition, the market price is also influenced by a random shock. Thus, the market price may be low due to either (1) overproduction on the part of the firms or (2) bad luck. However, since individual firm quantity choices are unobservable to all firms, it is impossible to verify the true cause of the low market price. This typically means that the firms would be unable to sustain a collusive agreement. However, using a punishment system where the firms punish for x periods if the market price ever drops below some level $\underline{p}$ regardless of the reason (either bad luck or overproduction) the firms are able to sustain a noncooperative collusive agreement.

[5] Green and Porter (1984), *Econometrica*, Noncooperative Collusion Under Imperfect Price Information, vol. 52:1, 87-100.