4. **Mixed Strategy Nash Equilibrium**

Solving for a mixed strategy Nash equilibrium means that at least one of the players prefers not to pick one action with certainty, but rather wants to be unpredictable (the *matching pennies* game presented in class is a good example, as is poker or a penalty kick in soccer). The question that we thus have to answer here is with what probability the Row player (player 1) chooses the options available to her and with what probability the Column player (player 2) chooses either C or D.

What we are looking for more concretely is a pattern where both players are indifferent between playing either C or D, given what the other person does. Let's see what player 1 can do to make player 2 indifferent between her strategies:

Let $p$ be the probability that player 1 plays $C_1$ and $(1-p)$ the probability that player 1 plays $D_1$. We now want to ensure that player 2 is indifferent between playing $C_2$ or $D_2$, i.e.,

$$\begin{aligned}
U_2(C_2) &= U_2(D_2) \\
p \times U_2(C_2|C_1) + (1-p) \times U_2(C_2|D_1) &= p \times U_2(D_2|C_1) + (1-p) \times U_2(D_2|D_1) \\
p \times (-3) + (1-p) \times 2 &= p \times 2 + (1-p) \times (-3) \\
-5p + 2 &= 5p - 3 \\
5 &= 10p \\
p &= 0.5
\end{aligned}$$

Player 2's utility for $\{\frac{1}{2}C_1, \frac{1}{2}D_1\}$ is $\frac{1}{2}(-3) + \frac{1}{2}(2) = -1.5 + 1 = -0.5$ and $\frac{1}{2}(2) + \frac{1}{2}(-3) = 1 + (-1.5) = -0.5$. Thus, player 2 is indifferent between her options when player 1 plays C and D half of the time each; in both cases the expected payoff is $-0.5$.

The same thought process and calculation applies to player 2's behavior as well, and she will also play C and D half of the time each (feel free to go through the above steps again). The strategy pair

Row player: $\{\frac{1}{2}C_1, \frac{1}{2}D_1\}$,
Column player: $\{\frac{1}{2}C_2, \frac{1}{2}D_2\}$

is therefore the mixed strategy Nash equilibrium that we are looking for. No deviation from this pattern pays off, and choosing either C or D half of the time is a best response to what the other player is doing.

5. **Repeated games**

This is what the game looks like in strategic form:

| | C | D |
|---|---|---|
| C | 5,5 | 0,9 |
| D | 9,0 | 3,3 |

(a) Find the Nash equilibrium of the stage game (a single-play of PD).

The Nash equilibrium for the stage game is (D,D). This is the only outcome in which no player, by changing her part of the strategy profile *unilaterally*, can improve her payoff. In other words, there is no incentive to switch if the other person stays put.

(b) Find the discount factor necessary to prevent a player from defecting if both players play grim trigger strategies: play C in the first period; always continue to play C if the other player does likewise, and always defect if the other player has ever defected.

Let's look at the sequence of moves that we should observe here (the indices denote the round of play; player 1 is the row player):

$$\text{Player 1: } C_1 \;\; C_2 \;\; C_3 ...$$
$$\text{Player 2: } C_1 \;\; C_2 \;\; C_3 ...$$

If both players play grim trigger, they cooperate indefinitely. However, there might be a temptation to "cheat" at some point in time, i.e., to switch from grim trigger to always defecting from some point in time $L$ on, in order to secure a higher overall payoff. What we are interested in in this case is a pattern such as this (say player 1 deviates):

$$\text{Player 1: } C_1 \;\; C_2 \;\; C_3 \;\; ... \;\; C_{L-1} \;\; D_L \;\; D_{L+1} \;\; D_{L+2} \;\; ...$$
$$\text{Player 2: } C_1 \;\; C_2 \;\; C_3 \;\; ... \;\; C_{L-1} \;\; C_L \;\; D_{L+1} \;\; D_{L+2} \;\; ...$$

If there was no discounting, cheating would never pay off, as the player deviating would lose in the long run despite a short-term reward of 9 vs. the usual 5, as in all future rounds the payoffs would be only 3 as opposed to 5 before. Yet, if the players value the present more than the future, receiving a payoff of 9 now might be worth a lot to them when compared to the loss of 2 units they incur in each future round as a consequence of their defection and the other player's punishment of this behavior. How highly do they have to value the present relative to the future to make a defection profitable, though?

We know that under cooperation, each player receives a payoff of

$$U(Grim|Grim) = 5 + \delta \times 5 + \delta^2 \times 5 + \delta^3 \times 5 + ... = \frac{5}{1-\delta}$$

Contrast this to the payoff, called $d$ here, a player gets from switching to defection at some time $L$; $U(Grim + Defect|Grim)$ is given by:

$$5 + \delta \times 5 + \delta^2 \times 5 + ... + \delta^{L-1} \times 5 + \delta^L \times 9 + \delta^{L+1} \times 3 + \delta^{L+2} \times 3 + ... = d$$

We are now looking for a potential value of $\delta$ such that $d > \frac{5}{1-\delta}$. As we know that the payoffs in the first $L-1$ rounds are the same, we can safely disregard these and concentrate on what happens from the period of defection, $L$, on. We are thus looking for a value of $\delta$, s.t.,

$$\delta^L \times 9 + \delta^{L+1} \times 3 + \delta^{L+2} \times 3 + ... \geq \frac{5\delta^L}{1-\delta}$$

Dividing by $\delta^L$ throughout and rewriting 9 as (6+3), we get

$$\begin{aligned}
6 + 3 + \delta^1 \times 3 + \delta^2 \times 3 + ... &\geq \frac{5}{1-\delta} \\
6 + \frac{3}{1-\delta} &\geq \frac{5}{1-\delta} \\
6 \times (1-\delta) + 3 &\geq 5 \\
6 \times (1-\delta) &\geq 2 \\
6 - 6 \times \delta &\geq 2 \\
-6 \times \delta &\geq -4 \\
\delta &\leq \frac{2}{3}
\end{aligned}$$

This is the solution we have been looking for. If $\delta$ is smaller than 2/3, then it pays to defect at some point in time. For a discount factor $\delta$ greater than 2/3, however, defection hurts you overall and you'd therefore stick with cooperation at all times.[1]

---

[1] When $\delta = \frac{2}{3}$, the player is indifferent between either cooperating forever or defecting at some point, and the discounted payoffs are 15 in both cases.

(c) Does this pair of strategies form a Nash equilibrium? Does it form a subgame perfect equilibrium (SPE)? Why (or why not)?

If both players are playing grim trigger, then it depends on the discount factor $\delta$ whether we have a Nash equilibrium or not. If $\delta = 0.5$, for example, there is a unilateral incentive to defect even though the other person plays grim trigger (if you defect in the first round, you secure yourself a payoff of 12 that way, as opposed to a payoff of 10 which you'd get for always cooperating, i.e., sticking to grim trigger). Thus, whenever $\delta < 2/3$, grim trigger is not a Nash equilibrium and therefore also not a subgame perfect equilibrium.

In the case where $\delta \geq 2/3$, grim trigger is both a Nash equilibrium and subgame perfect. If you value the future highly, defection doesn't pay, and we thus have (C,C) as a Nash equilibrium. Yet, for subgame perfection we have to make sure that even in those states of the game which we will never reach the strategy the players pursue is still optimal (you may think of this as an insurance against someone making a mistake at some point).

Note that no matter what you do, after there has been a defection once, the other player will always defect and there is thus no way to ever get back to the (C,C) state. You might switch back to C, yet this would give you a lower payoff than D, and thus when you are in a state where the other player defects, it is your best response to also defect at all times. Grim trigger is thus subgame perfect (recall that even in states which you can't reach as long as the players act as their strategy tells them to, such as the $D$ case here, your strategy must still be a best response for subgame perfection to hold).

(d) Now find the discount factor necessary to prevent a player from defecting if both players play tit-for-tat strategies: play C in the first period; always play whatever the other player played in the previous period thereafter.

The reasoning here is parallel to that in part (b) of this question; with tit-for-tat we get a different pattern of behavior, though. First, let's once more establish the benchmark against which we measure the profitability of a defection. As before, when both players stick to tit-for-tat (and thus cooperate indefinitely), we have the following payoff:

$$U(TFT|TFT) = 5 + \delta \times 5 + \delta^2 \times 5 + \delta^3 \times 5 + ... = \frac{5}{1-\delta}$$

Now, let's say that there is a defection by player 1 at some point $L$:

$$\begin{array}{l} \text{Player 1: } C_1 \;\; C_2 \;\; ... \;\; C_{L-1} \;\; D_L \;\; C_{L+1} \;\; D_{L+2} \;\; C_{L+3} \;\; ... \\ \text{Player 2: } C_1 \;\; C_2 \;\; ... \;\; C_{L-1} \;\; C_L \;\; D_{L+1} \;\; C_{L+2} \;\; D_{L+3} \;\; ... \end{array}$$

From period $L$ on we thus get a pattern where one player defects while the other cooperates, and the next period it's the other way round.[2]

As before, we can disregard the first $L-1$ periods in our utility calculations, as they are unaffected by the eventual defection in period L. The payoffs from period L on are:

$$\begin{aligned} TFT\ with\ defection &= 9 \times \delta^L + 0 \times \delta^{L+1} + 9 \times \delta^{L+2} + ... \\ TFT &= 5 \times \delta^L + 5 \times \delta^{L+1} + 5 \times \delta^{L+2} + ... \end{aligned}$$

Note that we can rewrite this as

$$\begin{gathered} 5 \times \delta^L + 5 \times \delta^{L+1} + 5 \times \delta^{L+2} + 5 \times \delta^{L+3} + ... = \\ 5 \times \delta^L + 5 \times \delta^{L+1} + \delta^2[5 \times \delta^L + 5 \times \delta^{L+1}] + ... \end{gathered}$$

This simplifies our life quite a bit, as we now only have to find the value of $\delta$ for which

$$\begin{aligned} 9 \times \delta + 0 \times \delta^2 &\geq 5 \times \delta + 5 \times \delta^2 \\ 4\delta &\geq 5\delta^2 \\ 4 &\geq 5\delta \\ 0.8 &\geq \delta \end{aligned}$$

Hence, if $\delta$ is smaller than 0.8, a defection pays off. As in (b), if $\delta = 0.8$, the players are equally well off if they defect once and keep playing tit-for-tat subsequently or if they keep on cooperating forever. For all $\underline{\delta \text{ greater than } 0.8}$, a defection does not pay off overall, as the losses that are incurred in subsequent rounds outweigh the gains that are made in the short run.

---

[2] We have seen a different variety of TFT in class, where the defector cooperates twice after the defection and thus establishes mutual cooperation again. The question in this problem set does not mention "play C twice after a defection," though.

(e) Does this pair of strategies form a Nash equilibrium? Does it form a subgame perfect equilibrium? Why (or why not)?

Tit-for-tat is a Nash equilibrium for $\delta \geq 0.8$, yet it is not subgame perfect. When $\delta < 0.8$, tit-fot-tat is not even a Nash equilibrium – we have seen in part (d) that it then pays off to deviate from TFT –, and therefore by definition also not subgame perfect.[3]

Let's look at the case of $\delta \geq 0.8$ in some more detail. We have seen before that when we value the future enough, then it does not pay to defect at some point $L$ and thereby trigger a chain of (C,D), (D,C), (C,D,), ... decisions. This is sufficient for a Nash equilibrium: it does not pay off to deviate unilaterally from the chosen strategy.

By contrast, TFT is not subgame perfect for $\delta \geq 0.8$. We can check whether some other strategy gives better results after an initial defection, and I propose that cooperating twice after a defection does so (we have already encountered this strategy in the lectures). What I have to show is that in the two periods L+1 and L+2 combined my payoffs are higher when I cooperate twice than when I defect in L+2, i.e., stick to my TFT strategy.

We know from (d) that (C,C) from period $L+3$ on is better than jumping between (C,D) and (D,C) all the time. I thus have to look only at the periods L+1 and L+2 and compare the payoffs there. If I find that cooperating twice after my initial defection gives me a higher payoff than cooperating once and then defecting, I have shown that TFT is not a best repsonse:

$$\begin{aligned} 9 \times \delta^{L+1} + 0 \times \delta^{L+2} &\leq 5 \times \delta L + 1 + 5 \times \delta^{L+2} \\ 9 &\geq 5 + 5 \times \delta \\ 4 &\geq 5 \times \delta \\ 0.8 &\geq \delta \end{aligned}$$

This is of course only true for $\delta = 0.8$. As soon as we choose a $\delta$ such that $0.8 < \delta \leq 1$, this inequality does not hold any more and the payoff from cooperating twice is higher than that which would be obtained by sticking to TFT. Thus, TFT is not a best response after an initial defection and thus not subgame perfect.

---

[3] Recall that all subgame perfect equilibria must also be Nash equilibria, but not all Nash equilibria are subgame perfect — as is the case with tit-for-tat here.