22.- Repeated Games and Reputation

- In the real world, agents interact with each other repeatedly over time instead of just once in a static setting.
- If players interact repeatedly over time, future behavior can be affected by current behavior. Therefore, short-term actions must be balanced against their possible long-term repercussions.
- When players interact with each other repeatedly, players’ **reputations** can become a key determinant of their behavior.
- As a result, in repeated games we may observe behavior that could not be sustained if the game were played only once. An example will be: Cooperation in the prisoners’ dilemma.

- **A repeated game:** Is played over discrete periods of time. We let t denote any given period in the game. We let T denote the total number of periods in which the game is played (with $T = \infty$ as a possibility).
- In each period $t = 1, \ldots, T$, players play a static **stage game** where they simultaneously and independently select actions. The action space of the stage game is:

$$A = A_1 \times A_2 \times \cdots \times A_n$$

- Where A_i denotes the action space of player i, with $i = 1, \ldots, n$.
- **Payoff functions for the stage game** are denoted as u_i, where $u_i(a)$ is the stage game payoff for player i if action profile $a = (a_1, a_2, \ldots, a_n)$ is played.

- We assume that at any period t, **the entire history of play** from periods $1, \dots, t-1$ **is observed by every player**.
- **Overall payoffs for the entire game**: Are given by the **discounted sum of the stage game payoffs** for periods $t = 1, \dots, T$.
- As we have done in sequential games, we will focus on studying **subgame perfect equilibria** (SPE) in repeated games.
- We will begin by studying two-period repeated games. That is, a repeated game where $T = 2$.

- **A Two-Period Repeated Game:** Suppose the following stage game is played by two players in periods $t = 1,2$:

1 \ 2	X	Y	Z
A	4, 3	0, 0	1, 4
B	0, 0	2, 1	0, 0

- Focus on pure strategies. This repeated game has 36 total possible outcomes (6 in each period). Therefore, its extensive form representation is complicated (this is typically the case in repeated games).

- Suppose for simplicity that both players are perfectly patient between both time periods, so both of their discount factors are $\delta = 1$.

- Therefore, the overall payoffs for the entire game are simply given by the sum of the payoffs in each time period.

- This game has only one proper subgame: The stage game played in period t=2.

- How do the continuation payoffs for the subgame look like? We can compute them simply by adding the payoffs of the outcome in the first period to the stage-game payoff matrix…

- Continuation payoffs for subgame in t=2:
- If (A,X) is played in t=1:

1 \ 2	X	Y	Z
A	8, 6	4, 3	5, 7
B	4, 3	6, 4	4, 3

- If (A,Y) is played in t=1:

1 \ 2	X	Y	Z
A	4, 3	0, 0	1, 4
B	0, 0	2, 1	0, 0

- Continuation payoffs for subgame in t=2:
- If (A,Z) is played in t=1:

1 \ 2	X	Y	Z
A	5, 7	1, 4	2, 8
B	1, 4	3, 5	1, 4

- If (B,X) is played in t=1:

1 \ 2	X	Y	Z
A	4, 3	0, 0	1, 4
B	0, 0	2, 1	0, 0

- Continuation payoffs for subgame in t=2:
- If (B,Y) is played in t=1:

1 \ 2	X	Y	Z
A	6, 4	2, 1	3, 5
B	2, 1	4, 2	2, 1

- If (B,Z) is played in t=1:

1 \ 2	X	Y	Z
A	4, 3	0, 0	1, 4
B	0, 0	2, 1	0, 0

- **Subgame perfect equilibria:** Let us focus on pure-strategies. We know that in any SPE, players MUST play a Nash equilibrium in the subgame played in t=2.
- We have described all the possible continuation payoffs for this subgame. It is easy to verify that in all cases, the subgame has the same Nash equilibria as the stage game. Namely: $(\boldsymbol{A}, \boldsymbol{Z})$ and $(\boldsymbol{B}, \boldsymbol{Y})$.
- Thus, **in any SPE of this game, either (A, Z) or (B, Y) must be played in $t = 2$.**

- What about in period t=1? What strategies can we observe in t=1 in an SPE?
- First, it is fairly obvious to see that playing a Nash equilibrium of the stage game in period t=1 and then a Nash equilibrium of the stage game in period t=2 constitutes an SPE.
- In fact we have the following general result:

Result: Take ANY repeated game. Then any sequence of stage Nash equilibrium profiles can be supported as the outcome of a subgame perfect Nash equilibrium.

- Thus, having either (A,Z) or (B,Y) in period t=1 followed by either (A,Z) or (B,Y) in period t=2 can be supported as the outcome of SPE in this example.

- The previous result is neither surprising, nor very interesting. What is interesting is the following question: **Can we support non-equilibrium outcomes in period t=1 in a SPE of this game?**

- The answer is **YES, as long as the equilibrium selected in t=2 is determined by the players' actions in period t=1.**

- For example: Can we make the players play the outcome (A,X) in period t=1 in a SPE?
- Note that (A,X) is NOT a Nash equilibrium of the stage game.
- Also note that in any SPE, players must either play (A,Z) or (B,Y) in period t=2.
- Who has the incentive to deviate from (A,X) in period t=1? Player 1’s best response to “X” is indeed “A”, so player 1 does not have an incentive to deviate.
- On the other hand, player 2’s best response to “A” is “Z”. Thus, player 2 is the only one with an incentive to deviate from (A,Z) in period t=1.

- Thus, if (A,X) is to be played in t=1 in any SPE, it must be the case that the Nash equilibrium selected in period t=2 is such that it **rewards player 2 if he adheres to (A,X) in t=1, and it punishes player 2 if he deviates from (A,X) in t=1**.
- The stage-game payoff to player 2 under equilibrium (B,Y) are 1, and they are 4 under equilibrium (A,Z). Therefore, player 2 must be rewarded with equilibrium (A,Z) if he adheres to (A,X) in t=1, and he must be punished with (B,Y) if he deviates.
- The question is: Is the punishment/reward strong enough to make player 2 adhere to (A,X) in t=1?

- If player 2 adheres to (A,X) in t=1, his overall payoffs would be:

$$\mathbf{3+4=7}$$

- On the other hand, if player 2 deviates in t=1, his best-response would be to choose "Z", which would earn him a payoff of 4, followed by the punishment (B,Y) which earns him a payoff of 1. Therefore his overall payoff from deviating would be:

$$\mathbf{4+1=5}$$

- Therefore, player 2 is better off adhering to (A,X) in t=1. Therefore, **observing the outcome (A,X) is consistent with a SPE in this game**. The key is that the punishment/rewards in the second stage are **<u>credible</u>** because they are Nash equilibria of the stage game.
- Note that (A,X) would never be observed in equilibrium if the game were played only once (instead of twice).

- Note: We assumed that player 2 is perfectly patient, so that his discount rate is $\delta = 1$. What if he is not perfectly patient? Can we still have (A,X) in t=1 as part of a SPE?
- It depends on the degree of impatience of player 2.
- If he adheres to (A,X) in t=1, his overall payoffs would be:

$$\mathbf{3 + \boldsymbol{\delta} \cdot 4}$$

- If he deviates, his overall payoffs would be

$$\mathbf{4 + \boldsymbol{\delta} \cdot 1}$$

- Thus, (A,X) can be observed in t=1 in an SPE if and only if

$$\mathbf{3 + \boldsymbol{\delta} \cdot 4 \geq 4 + \boldsymbol{\delta} \cdot 1}$$

- That is, if and only if

$$\boldsymbol{\delta} \geq \mathbf{\frac{1}{3}}$$

- This measures the degree of "patience" that player 2 must have. If he is "too impatient" (in this case, if $\delta < \frac{1}{3}$) then (A,X) cannot be observed in an SPE.

- **Example:** Consider the prisoner's dilemma game we have studied before:

1 \ 2	**C**	**D**
C	2, 2	0, 3
D	3, 0	1, 1

- Suppose this game is played **twice**. Is it possible to observe cooperation in t=1 in an SPE?
- **Note first that this stage game has a UNIQUE Nash equilibrium (D,D).** Therefore in any SPE the only outcome that can be observed in t=2 is (D,D). Therefore we cannot have different "rewards" and "punishments" in t=2.

- Payoffs are symmetric to both players. Each player has the incentive to deviate from (C,C) in t=1 by choosing “D”. Suppose they both have the same discount factor δ The overall payoff each player would obtain if they deviate from (C,C) in t=1 is:

$$3 + \delta \cdot 1$$

- On the other hand, if they adhere to (C,C) in t=1, their payoff would be:

$$2 + \delta \cdot 1$$

- Thus, in order to have cooperation in t=1 we must have $2 + \delta \cdot 1 \geq 3 + \delta \cdot 1$. Clearly this is impossible for any value of δ. We conclude that **if this game is played only twice, cooperation can never be observed in any period in an SPE. We will see below that this result changes if the game is repeated many more times (instead of only twice).**

- **An infinitely repeated game:** Suppose $T = \infty$, so the stage game keeps being played indefinitely.
- To make sure that the present discounted value of payoffs is a well defined quantity in infinitely repeated games, we will maintain the assumption that discount factors are strictly less than one. That is, $\delta_i < 1$ for each $i = 1, \dots, n$.
- Consider the sum:

$$v \equiv 1 + \delta + \delta^2 + \delta^3 + \delta^4 + \cdots$$

- If $|\delta| < 1$, then the sum v is well defined (finite) and it can be simplified into a very sort expression. Note that

$$\boldsymbol{v \equiv 1 + \delta \cdot \left[1 + \delta + \delta^2 + \delta^3 + \delta^4 + \cdots\right] = 1 + \delta \cdot v}$$

- Therefore if $|\delta| < 1$, we have:

$$\boldsymbol{v = 1 + \delta \cdot v}$$

- Therefore if $|\delta| < 1$,

$$\boldsymbol{v = \frac{1}{1 - \delta}}$$

- We will use this formula in what follows…
- **SPE in infinitely repeated games:** Immediately we can identify a particular type of SPE:

Result: Take ANY infinitely repeated game. Then any sequence of stage Nash equilibrium profiles can be supported as the outcome of a subgame perfect Nash equilibrium.

- Again, this result is not very interesting. **What we want to find out is whether outcomes that are not Nash equilibria in the stage game can be sustained in an SPE of an infinitely repeated game**.
- In an infinitely repeated game, there are infinitely many possible types of strategies and histories of play. However, we focus on a specific type of strategy which we refer to as **trigger strategies.**
- **Trigger strategies:** Prespecify two types of action profiles in the stage game:
  - **A "cooperative profile".**
  - **A "punishment profile".**

- Trigger strategies have two types of action profiles for the stage game:
- **Cooperative profile:** Label it as c_t. We allow it to change with t, although in many examples we will assume it is the same for all t.
- **Punishment profile:** Label it as p**.** We will assume that the punishment profile is the same for all t. **Credible punishments** are always Nash equilibrium profiles of the stage game. Thus, the **punishment profile p will always be a Nash equilibrium profile of the stage game**.
- Trigger strategies specify the following type of behavior:

"Play cooperative profile c_t in period t as long as everybody has played the cooperative profile in periods $1, \ldots, t-1$. Otherwise, if somebody deviated from the cooperative profile in the past, play the punishment profile p for ever."

- Under what conditions will a trigger strategy constitute an SPE?
- Consider any period t. If player i adheres to the cooperative profile, then his continuation payoffs for the game will be:

$$\boldsymbol{u_i(c_t) + \delta_i \cdot u_i(c_{t+1}) + \delta_i^{\,2} \cdot u_i(c_{t+2}) + \delta_i^{\,3} \cdot u_i(c_{t+3}) + \cdots}$$

- What about if player i deviates from the cooperative profile in period t?
- Let $\boldsymbol{d_{it}}$ denote the profile where all players except i play the cooperative profile c_t but player i deviates and plays the best response he has against c_t.

- Then the continuation payoffs for i from deviating from the cooperative profile are:

$$\boldsymbol{u_i(d_{it}) + \delta_i \cdot u_i(p) + \delta_i{}^2 \cdot u_i(p) + \delta_i{}^3 \cdot u_i(p) + \cdots}$$

- Thus, this trigger strategy will be an SPE if, at every period t and each player $i = 1, \dots, n$, we have:

$$\boldsymbol{u_i(c_t) + \delta_i \cdot u_i(c_{t+1}) + \delta_i{}^2 \cdot u_i(c_{t+2}) + \delta_i{}^3 \cdot u_i(c_{t+3}) + \cdots}$$
$$\geq$$
$$\boldsymbol{u_i(d_{it}) + \delta_i \cdot u_i(p) + \delta_i{}^2 \cdot u_i(p) + \delta_i{}^3 \cdot u_i(p) + \cdots}$$

- **Let us focus on trigger strategies where the cooperation profile is the same for all periods t. Label this profile simply as c.** Then, this strategy will be an SPE of the game if, for each player $i = 1, \dots, n$:

$$u_i(c) + \delta_i \cdot u_i(c) + {\delta_i}^2 \cdot u_i(c) + {\delta_i}^3 \cdot u_i(c) + \cdots$$
$$\geq$$
$$u_i(d_i) + \delta_i \cdot u_i(p) + {\delta_i}^2 \cdot u_i(p) + {\delta_i}^3 \cdot u_i(p) + \cdots$$

- Where d_i is the strategy profile where everybody but player i cooperates while i deviates by playing his best response.

- The previous equation can be simplified to:

$$u_i(c) + u_i(c) \cdot \delta_i \cdot \left[1 + \delta_i + {\delta_i}^2 + \cdots\right]$$
$$\geq$$
$$u_i(d_i) + u_i(p) \cdot \delta_i \cdot \left[1 + \delta_i + {\delta_i}^2 + \cdots\right]$$

- Using our previous results about infinite sums, this can be further simplified to:

$$u_i(c) + \frac{u_i(c) \cdot \delta_i}{1 - \delta_i} \geq u_i(d_i) + \frac{u_i(p) \cdot \delta_i}{1 - \delta_i}$$

- Rearranging the previous inequality, it becomes:

$$\delta_i \geq \frac{u_i(d_i) - u_i(c)}{u_i(d_i) - u_i(p)}$$

- Therefore, the trigger strategy we described is an SPE if and only if, for each player $i = 1, \ldots, n$:

$$\boldsymbol{\delta_i \geq \frac{u_i(d_i) - u_i(c)}{u_i(d_i) - u_i(p)}}$$

- Intuitively: All players must be "sufficiently patient" where "sufficiently patient" is described by the threshold given above for the discount factor δ_i.

- **Example:** Consider again the prisoner's dilemma game we have studied before:

1 \ 2	**C**	**D**
C	2, 2	0, 3
D	3, 0	1, 1

- Suppose this game is infinitely repeated. **Under what conditions can cooperation be sustained as an SPE?**
- **Note first that there is only one available punishment because this stage game has a unique Nash equilibrium profile: (D,D).**

- We have

$$u_i(c) = 2 \text{ for } i = 1,2.$$
$$u_i(d_i) = 3 \text{ for } i = 1,2.$$
$$u_i(p) = 1 \text{ for } i = 1,2.$$

- Using our previous formula, we know that cooperation can be sustained if and only if, for each $i = 1,2$:

$$\delta_i \geq \frac{3-2}{3-1} = \frac{1}{2}$$

- That is, if and only if the discount factor is at least equal to $\frac{1}{2}$ for each player.

- **Example:** Suppose instead that the payoffs in the prisoner's dillemma are:

1 \ 2	**C**	**D**
C	3, 4	0, 7
D	5, 0	1, 2

- Suppose this game is infinitely repeated. Under what conditions can cooperation be sustained as an SPE?
- Again, there is only one available punishment because this stage game has a unique Nash equilibrium profile: (D,D).

- We now have

$$u_1(c) = 3 \text{ and } u_2(c) = 4$$
$$u_1(d_1) = 5 \text{ and } u_2(d_2) = 7$$
$$u_1(p) = 1 \text{ and } u_2(p) = 2$$

- Using our previous formula, we know that cooperation can be sustained if and only if:

$$\delta_1 \geq \frac{5-3}{5-1} = \frac{1}{2} \quad \text{and} \quad \delta_2 \geq \frac{7-4}{7-2} = \frac{3}{5}$$

- That is, if and only if the discount factor is at least equal to $\delta_1 \geq \frac{1}{2}$ and $\delta_2 \geq \frac{3}{5}$.

- **A general result:** OK, so we have seen that outcomes that are not Nash equilibria in the stage game can be sustained in infinitely repeated games if players are sufficiently patient.
- In terms of payoffs, this means that payoffs that cannot be achieved in Nash equilibria of the stage game CAN be achieved and sustained in infinitely repeated games
- A more general question in infinitely repeated games is: What range of payoffs can be sustained in infinitely repeated games?
- This question leads to a very important result in game theory.

- We have the following result:

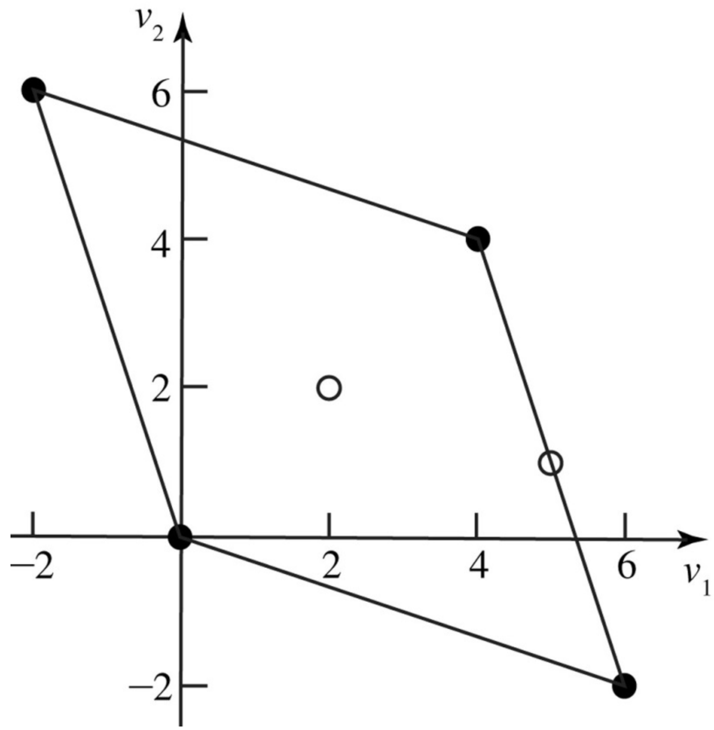
Result: If players are sufficiently patient (that is, if their discount factors δ_i are close enough to 1), then any weighted average of the payoffs in the stage game can be sustained in an infinitely repeated game.

- More precisely, for any weighted average of the payoffs in the stage game, we can always find a trigger strategy such that the SPE payoffs of the game correspond exactly to that weighted average.

- We can gain some intuition graphically. Consider the stage game:

1 \ 2	C	D
C	4, 4	-2, 6
D	6, -2	0, 0

- How does the set of all weighted averages of the stage-game payoffs look like?
- We can represent it graphically as follows…



- The above convex set whose vertices are given by the four possible stage-game payoffs constitutes the set of all possible payoff levels that can be achieved in an infinitely repeated version of this game if players are patient enough (if their discount factors are close enough to 1). The two empty circles display two examples of payoff levels that can be achieved.

- The proof of the previous result is mathematically involved and well beyond the scope of our course.
- Therefore we will only mention this result without proof and we will not get into the details.
- What is interesting about it is to realize once again how moving from a static game to a sequential game can vastly increase the range of possible outcomes that can be observed in equilibrium.