3.11 Repeated Games

In many situations players will interact on many occasions.

Suppose two players play a matrix game n times.

In such a case a player can condition the action he takes in the i-th realisation of the game on the history of the game up to that point, denoted H_{i-1}.

H_{i-1} can be described by the pairs of actions used by the players in the first $i-1$ rounds.

The Prisoner's Dilemma

Intuition: The police have arrested two gangsters, but only have enough evidence to convict them of tax evasion. The gangsters cannot communicate. It can thus be assumed that their decisions are made simultaneously.

During interrogation, the police say that if a gangster gives evidence against the other (D - defect), then he will receive a reduced sentence.

In only one gives evidence against the other, then he goes free and the other receives 20 years in jail for extortion.

If both give evidence, then they both receive 10 years for extortion.

If neither give evidence (both play C - cooperate with each other), then they both receive 2 years for tax evasion.

The Prisoner's Dilemma

If we assume that the payoff of each gangster is the number of years he spends free in the next 20 years, then the payoff matrix for this game is

	C	D
C	(18,18)	(0,20)
D	(20,0)	(10,10)

The Prisoner's Dilemma

Suppose the second gangster cooperates. The first gangster goes free if he defects, otherwise he gets two years in jail.

Suppose the second gangster defects. The first gangster gets 10 years if he defects, otherwise he gets 20 years in jail.

Hence, defection dominates cooperation. The unique Nash equilibrium is for both players (gangsters) to defect.

However, both players would obtain a greater payoff if they both cooperated.

In the general form of the prisoner's dilemma game, it is assumed that players obtain a greater payoff by always playing (C, C) than by alternating between (C, D) and (D, C).

The Prisoner's Dilemma

We have seen that the concept of correlated equilibria (using communication) can enable players to find a suitable solution to a co-ordination or anti-coordination game.

However, if there is only one pure Nash equilibrium of a game, then it is the only correlated equilibrium.

It follows that the concept of a correlated equilibrium cannot be used to ensure that the gangsters cooperate.

What systems can be used in games of the form of the prisoner's dilemma to ensure "cooperation"?

The Prisoner's Dilemma

One way of obtaining cooperation is by means of a contract.

Such a contract defines penalties to be paid when a player uses an inappropriate action. These penalties define an induced game, whose payoff matrix takes into account these penalties.

The penalties should be chosen to ensure that at a Nash equilibrium of the induced game both players cooperate.

In terms of the example given, cooperation can be achieved via threats from the godfather to punish those who "sing".

The Prisoner's Dilemma

Another possibility is that individuals play a game many times.

In such a case, players can use strategies such as tit-for-tat, i.e. start by cooperating and then use the action that the other player used in the previous round.

Two players using such a strategy will always cooperate. However, such players will react against defections by the other player and thus avoid paying the cost of being "a sucker" (cooperating when the other player defects).

Hence, the best response to an individual playing tit-for-tat may well be to also play tit-for-tat. In this case, tit-for-tat would be an equilibrium in the repeated prisoner's dilemma.

Definition of an n-Repeated Game

Suppose in the standard game G_1, player i can take an action from S_i and the payoff of player i when player 1 plays $A \in S_1$ and player 2 plays $B \in S_2$ is given by $R_i(A, B)$.

In the n-repeated version of G_1, denoted G_n, in each round j, $1 \leq j \leq n$, the set of actions available to the players and the payoffs received are given as above.

The total payoff of a player in the repeated game is simply the sum of the payoffs he/she receives in each round.

The action taken by a player in round i can depend on the history H_{i-1}. We normally only consider pure strategies, i.e. given H_{i-1} each player picks a particular action with probability 1.

The 2-Repeated Prisoner's Dilemma

In such a game, the action of a player in round 2 can depend on the pair of actions played in round 1.

Since a player's strategy defines what action he/she takes in round 1, it suffices to condition the action taken in round 2 on the action taken by the other player in round 1.

The (partial) strategy of the players in round 1 is defined by the action taken (2 possibilities). In round 2 for each of the 2 actions possibly taken by the other player in round 1, a player can take one of two actions (4 possibilities in total).

It follows that each player has 8 possible strategies. These are listed on the next page.

The 2-Repeated Prisoner's Dilemma

1. (C, [C/C, C/D]) - always cooperate.
2. (C, [C/C, D/D]) - tit-for-tat.
3. (C, [D/C, C/D])
4. (C, [D/C, D/D]) - cooperate and then defect.
5. (D, [C/C, C/D]) - defect and then cooperate.
6. (D, [C/C, D/D]) - pessimistic tit-for-tat.
7. (D, [D/C, C/D])
8. (D, [D/C, D/D]) - always defect.

The 2-Repeated Prisoner's Dilemma

We can define the payoff matrix for such a game, by considering the action pair used in each round of the game given the strategies used by the players.

Suppose the prisoner's dilemma game defined above is repeated twice and Player 1 plays tit-for-tat (Strategy 2) and Player 2 plays always defect (Strategy 8).

The action pairs played are (C, D) and (D, D) in round 1 and round 2, respectively.

Player 1 obtains 0 in round 1 and 10 in round 2. Her total payoff is 10.

Player 2 obtains 20 in round 1 and 10 in round 2. His total payoff is 30.

Matrix form of the 2-Repeated Prisoner's Dilemma

	1	2	3	4	5	6	7	8
1	(36,36)	(36,36)	(18,38)	(18,38)	(18,38)	(18,38)	(0,40)	(0,40)
2	(36,36)	(36,36)	(18,38)	(18,38)	(20,20)	(20,20)	(10,30)	(10,30)
3	(38,18)	(38,18)	(28,28)	(28,28)	(10,30)	(10,30)	(0,40)	(0,40)
4	(38,18)	(38,18)	(28,28)	(28,28)	(20,20)	(20,20)	(10,30)	(10,30)
5	(38,18)	(20,20)	(30,10)	(20,20)	(28,28)	(10,30)	(28,28)	(10,30)
6	(38,18)	(20,20)	(30,10)	(20,20)	(30,10)	(20,20)	(30,10)	(20,20)
7	(40,0)	(30,10)	(40,10)	(30,10)	(28,28)	(10,30)	(28,28)	(10,30)
8	(40,0)	(30,10)	(40,0)	(30,10)	(30,10)	(20,20)	(30,10)	(20,20)

The 2-Repeated Prisoner's Dilemma

In each round a player can ensure a reward of 10 by defecting. Hence, at a Nash equilibrium of the 2-repeated game, both players must obtain a reward of at least 20.

Using this fact to aid an exhaustive search for a pure equilibrium, the only pure Nash equilibria of the game given above are $(6, 6)$, $(6, 8)$, $(8, 8)$ and $(8, 8)$. At any of these equilibria both players defect in both periods.

Comparing the payoffs Player 1 obtains when he plays 6 and when he plays 8, it can be seen that strategy 8 (always defect) dominates strategy 6. It follows that $(8, 8)$, i.e. both players always defect is the only subgame perfect Nash equilibrium strategy in this game.

Note that strategy 6 is a pessimistic version of tit-for-tat. It starts out by defecting and then takes the same action as the opponent played in the previous round.

Equilibria in Repeated Games

Theorem: Any sequence of action pairs which are all Nash equilibria in the one-shot game G_1 is a Nash equilibrium in the repeated game G_n.

Proof: Suppose one of the players wishes to unilaterally change his/her action in one or more rounds.

If the original action pairs were all Nash equilibria, the payoff he/she obtains in those rounds is not greater than under the sequence of Nash equilibrium action pairs.

Since the payoff of a player in the repeated game is the sum of the payoffs in the individual rounds, it follows that such an individual cannot obtain a greater payoff in the repeated game by unilateral defection. Hence, a sequence of Nash equilibrium action pairs is a Nash equilibrium in the repeated game.

Equilibria in the n-repeated prisoner's dilemma

Theorem: At any Nash equilibrium of the n-repeated prisoner's dilemma (n finite), both players always defect.

Sketch Proof: This is by recursion. In the final round a player maximises his/her total payoff (given the payoff he/she has already obtained) simply by maximising the reward from the final round.

Whatever the opponent plays, a player maximises his/her reward by defecting. Hence, both players should defect in the final round.

Given this, the game reduces to an $n-1$-repeated game in which (as above) both players should defect in round $n-1$. Arguing by recursion, both players should defect in each round of the game.

Equilibria in the repeated prisoner's dilemma with a random number of rounds

If the number of rounds is known, then at any Nash equilibrium both players will always defect. Is there any form of repeated game in which tit-for-tat is a Nash equilibrium?

Normally individuals do not know how many times they will interact, although they may well be able to assess the likelihood of further interaction.

Consider the repeated prisoner's dilemma in which the probability that players meet again after any round is ω. Such a game is denoted G^{ω}

The number of rounds thus has a geometric distribution and the expected number of rounds is $\frac{1}{1-\omega}$. We expect tit-for-tat to be a "successful" strategy when the probability of meeting again, ω, is large

Equilibria in the repeated prisoner's dilemma with a random number of rounds

If one player plays "always defect", denoted D, then the best an opponent can do is also to always defect.

Hence, D is a Nash equilibrium in such a game.

At such an equilibrium both players obtain a payoff of 10 per round and the expected number of rounds is $\frac{1}{1-\omega}$.

The expected payoff of the players at such an equilibrium is thus $R_1(D, D) = \frac{10}{1-\omega}$.

Equilibria in the repeated prisoner's dilemma with a random number of rounds

When both players play tit-for-tat, denoted T they always cooperate.

Arguing as above, their expected reward is thus $R_1(T, T) = \frac{18}{1-\omega}$.

In order to check whether "tit-for-tat" is a Nash equilibrium, we first see how the following two strategies do against it:

1. Always defect.
2. Alternate between C and D.

Equilibria in the repeated prisoner's dilemma with a random number of rounds

Against tit-for-tat, a defector will obtain 20 in round 1. The game continues with probability ω.

Given the game continues, both players will always defect from round 1 onwards and the future expected reward of both is $\frac{10}{1-\omega}$ (starting at round 2 the game looks just like the repeated game in which both players always defect).

Hence, $R_1(D, T) = 20 + \frac{10\omega}{1-\omega}$.

Equilibria in the repeated prisoner's dilemma with a random number of rounds

Against tit-for-tat, a individual who alternates between D and C, denoted A, will obtain 20 in odd numbered rounds and 0 in even numbered rounds. The probability that round i is played is ω^{i-1}.

It follows that

$$R_1(A, T) = 20 + 20\omega^2 + 20\omega^4 + \ldots = \frac{20}{1-\omega^2}$$

Equilibria in the repeated prisoner's dilemma with a random number of rounds

Why does it suffice to just consider these two strategies?

It should be noted that if a T player plays with three opponents, one playing T, one playing D and one playing A, then all the 4 possible action pairs (C, C), (C, D), (D, C) and (D, D) are observed.

When any other player plays some strategy B against a T player, then $R_1(B, T)$ can be expressed as some linear combination of $R_1(T, T)$, $R_1(A, T)$ and $R_1(D, T)$, where the sum of the weights is 1.

It follows that if $R_1(D, T) \leq R_1(T, T)$ and $R_1(A, T) \leq R_1(T, T)$, then $R_1(B, T) \leq R_1(T, T)$ for any B, i.e. T is a Nash equilibrium.

Equilibria in the repeated prisoner's dilemma with a random number of rounds

It follows that T is a Nash equilibrium when

$$R_1(D,T) \leq R_1(T,T) \Rightarrow 20 + \frac{10\omega}{1-\omega} \leq \frac{18}{1-\omega}$$
$$R_1(A,T) \leq R_1(T,T) \Rightarrow \frac{20}{1-\omega^2} \leq \frac{18}{1-\omega}$$

Both these inequalities are satisfied if $\omega \geq 0.2$. It follows that tit-for-tat is a Nash equilibrium whenever players expect to meet at least $\frac{1}{1-0.2} = 1.25$ times in total.

Equilibria in Infinitely Repeated Games with a Discount

One may interpret the problem above as one in which the players meet an infinite number of times, but the reward obtained in round i is discounted by a factor ω^{i-1} in relation to the reward gained in round 1.

Using such an interpretation, the tit-for-tat strategy is a Nash equilibrium whenever future gains are valued high enough, i.e. ω is large.

In such cases, long term cooperation is beneficial compared to the short-term gains that might be obtained from defection.

"Stern" Strategies in Prisoner's Dilemma Type Games

The stern strategy S cooperates until the other player defects and then always defects.

After defecting against a stern player, it is clear that the defector should carry on defecting (since the stern player will always defect from then onwards and defection is the best response to this).

In order to check whether S is a Nash equilibrium, we only have to check that $R_1(S,S) \geq R_1(D,S)$.

"Stern" Strategies in Prisoner's Dilemma Type Games

It follows that if tit-for-tat is a Nash equilibrium, then "stern" is also a Nash equilibrium.

Note that in a population comprised purely of stern and tit-for-tat players, everyone will always cooperate.

One advantage of tit-for-tat occurs when with a small probability players make a mistake when choosing an action or perceiving the action taken by the other player.

If players use "stern", then after a defection cooperation breaks down. Using tit-for-tat (or a similar reactive but forgiving strategy), cooperation may well be re-established.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The repeated version of the Cournot game takes into account that firms will face a sequence of production decisions and take the previous behaviour of competitors into account.

What strategies in the repeated symmetric Cournot game would correspond to C, D and S in the repeated prisoner's dilemma game?

The corresponding strategy to C would be the strategy which when followed by both firms would maximise the sum of the profits (these would be split evenly between the firms).

The corresponding strategy to D would be the strategy both firms follow at the unique Nash equilibrium of the Cournot game.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The corresponding strategy to S would be the strategy: follow the profit maximisation strategy C as long as the other firm does the same, always play the strategy corresponding to the standard Cournot game after the other firm has deviated from the profit maximisation strategy.

For simplicity, we do not consider the analogous strategy to tit-for-tat.

Such defection only pays if the short-term reward from the initial defection outweighs the later loss from the lack of collusion. The only deviation from tit-for-tat we need to consider is the best response to the collusive strategy followed by repeating the Cournot equilibrium strategy.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

We assume that the payoff obtained in round i is discounted by a factor of ω^{i-1} compared to the payoff obtained in round 1 (as in the second interpretation of the iterated prisoner's dilemma game).

The method of finding the values of ω for which collusion is a Nash equilibrium is similar to the method used for the prisoner's dilemma.

1. We derive the symmetric action pair (i.e. both firms choose the same production level) that maximises the sum of profits. The corresponding action pair is (C, C).
2. We derive the Cournot equilibrium of the single-shot game and the associated profits. The corresponding action pair is (D, D),

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

3. We derive the best response to C and the payoffs obtained in this case. The corresponding action pair is (C, D).
4. We compare the discounted reward of a defector playing against a firm using S with the discounted reward of a firm playing S against another playing S.
5. Without loss of generality, we may assume that defection occurs in round 1.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

Consider the repeated Cournot game where the production of firm i in a period is x_i.

The price in a period is given by $p = 3 - \frac{x_1+x_2}{1000}$.

The costs incurred by firm i in a period are given by $c_i(x) = 100 + x_i$.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

Note that this is the problem considered in Section 3.6.

Player 1's best response to x_2 is $1000 - \frac{x_2}{2}$, when $x_2 \leq 2000$.

The Cournot equilibrium for this game is for both firm to produce $\frac{2000}{3}$ units per period.

The profits obtained by both players at the Cournot equilibrium are approximately 344.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

We need to derive the total production $x_{tot} = x_1 + x_2$ which maximises the total reward of the two firms.

The total costs incurred by the firms are
$c_1(x_1) + c_2(x_2) = 200 + x_1 + x_2 = 200 + x_{tot}$

The total revenue is $px_{tot} = (3 - \frac{x_{tot}}{1000})x_{tot} = 3x_{tot} - \frac{x_{tot}^2}{1000}$.

It follows that the total profit is
$R_{tot} = px_{tot} - c_1(x_1) - c_2(x_2) = 2x_{tot} - \frac{x_{tot}^2}{1000} - 200$.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

Differentiating with respect to x_{tot}, we obtain

$$\frac{\partial R_{tot}}{\partial x_{tot}} = 2 - \frac{x_{tot}}{500}.$$

It follows that the optimal joint production is 1000, i.e. each firm produces 500 units per period.

The total profit is 800, i.e. each firm makes a profit of 400 units at this "collusive solution".

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

We now derive the optimal response to this collusive level of production.

The analysis leading to the Cournot equilibrium shows that the optimal response of a firm when the other produces 500 units is to produce $1000 - \frac{500}{2} = 750$ units.

We now calculate the profit obtained by the "defector" in this case.

The price when the strategy pair is (500,750) is given by $p = 3 - \frac{1250}{1000} = 1.75$.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The "defector" obtains

$$R_2(500, 750) = 750 \times 1.75 - 100 - 750 = 462.5.$$

We can now find the values of ω for which collusion is a Nash equilibrium in the repeated game.

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The stern strategy, S, is to produce 500 units in a period as long as the other firm produces 500 units in a period and then always produce $\frac{2000}{3}$ units in a period.

A defector, denoted D, produces 750 units in the first period and then produces $\frac{2000}{3}$ units in a period.

Two firms playing S obtain 400 units per period. Their discounted reward is

$$R_1(S, S) = 400 + 400\omega + 400\omega^2 + \ldots = \frac{400}{1 - \omega}$$

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The payoff of a defector against a stern player is 462.5 in round 1 and 344 in subsequent rounds. Hence,

$$R_1(D, S) = 462.5 + 344\omega + 344\omega^2 + \ldots = 462.5 + \frac{344\omega}{1-\omega}.$$

Tit-for-tat is a Nash equilibrium in the repeated game if

$$\frac{400}{1-\omega} \geq 462.5 + \frac{344\omega}{1-\omega} \Rightarrow \omega \geq 0.5274.$$

Implicit Collusion in the Repeated Version of the Symmetric Cournot Game

The concept of collusion can be applied to asymmetric games. However, it is not so clear how to split the gains from collusion.

One possibility is that both firms split the gains from collusion equally between each other.

Another possibility is that the ratio between the profits of the firms at a collusive solution is equal to the ratio between the profits of the firms at the equilibrium of the corresponding one-shot Cournot (or Stackelberg) game.

These two possibilities are equivalent in the case of symmetric games.

Equilibria in infinitely Repeated Games Without Discounting

In the Hawk-Dove game playing the pair of actions (D, H) and (H, D) in alternate rounds is a Nash equilibrium in any n-repeated Hawk-Dove game, i.e. each pure Nash equilibrium pair is played half the time.

Compare this solution to the egalitarian correlated equilibrium "toss a coin" if the result is heads play (H, D) and (D, H).

If the game is repeated infinitely often, any solution where (H, D) is played a proportion p of the time and (D, H) is played a proportion $1 - p$ of the time is a Nash equilibrium.

Hence, any expected payoff vector at a correlated equilibrium which is a randomisation over the set of Nash equilibria can be achieved at a Nash equilibrium in an infinitely repeated game (in the sense of mean payoff per round).

Set of Possible Payoff Rates in Infinitely Repeated Games

Over an infinite horizon the mean (undiscounted) reward of the players per period is given by a weighted mean of all the payoff vectors over all the possible action pairs. The weights are the frequencies with which each action pair is played.

Hence, the set of attainable mean rewards is the same as the set of attainable expected rewards using a correlated equilibrium.

The Folk Theorem

The Folk Theorem: In an infinitely repeated game, for any mean payoff vector (per round) in which both players obtain at least their corresponding minimax reward there is a Nash equilibrium. There is no equilibrium at which at least one player obtains less than their minimax reward.

Such a Nash equilibrium is defined by a pair of strategies that give the required mean payoff vector plus the threat that if the other player "defects" from this strategy pair, then a player will play so as to minimise the payoff of the defector.

The Folk Theorem

It follows that in such an infinitely repeated game a Nash equilibrium which satisfies any sensible optimality condition must be Pareto optimal.

This follows since if a solution is not Pareto optimal, we can always find a Nash equilibrium at which one player obtains a greater mean reward without reducing the mean reward of the other player.

Example

Consider the infinitely repeated version of the chicken game

	A	B
A	(6,6)	(2,8)
B	(8,2)	(0,0)

Example

First we derive the minimax payoffs in the one-shot game (note that the game is symmetric).

Suppose Player 1 takes action A with probability p and Player 2 takes action A with probability q.

$$R_1(M_1, M_2) = 6pq + 2p(1-q) + 8(1-p)q = 2p + 8q - 4pq.$$

Example

We wish to find the strategy of Player 2 which minimises Player 1's expected reward. Hence, we calculate the derivative of Player 1's reward with respect to q (the decision of Player 2).

$$\frac{\partial R_1(M_1, M_2)}{\partial q} = 8 - 4p.$$

This is positive for all p. Hence, to minimise Player 1's payoff, Player 2 should choose q to be as small as possible.

It follows that Player 2 minimises Player 1's reward by always taking action B.

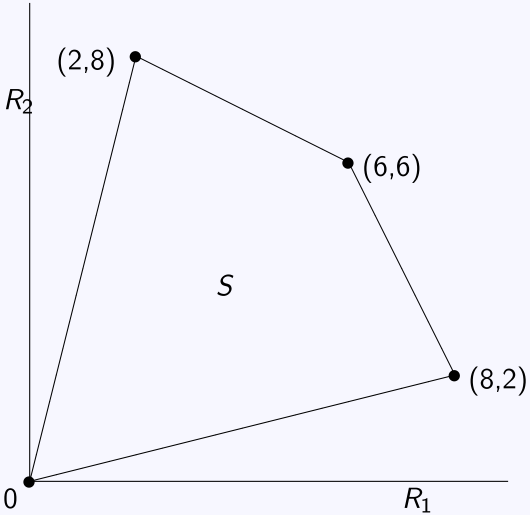
Example

If Player 2 plays B, then Player 1 should play A.

It follows that Player 1's minimax reward is 2. By symmetry, this is the minimax payoff of Player 2.

Hence, at any Nash equilibrium of the infinitely repeated game, both players must obtain a payoff of at least 2.

The Set of Attainable Mean Payoff Vectors



Example

Suppose we wish to find the Nash equilibrium which a) maximises the mean payoff of Player 1, b) maximises the sum of the mean payoffs.

a) The mean payoff vector at such a solution is the Pareto optimal mean payoff vector which gives Player 2 at least 2 and maximises the mean payoff of Player 1.

This is the mean payoff vector $(8, 2)$. The players have to play the action pair (B, A) in each round at such an equilibrium.

At such an equilibrium the threat, that a player will play B if the other deviates from this action pair, is essentially meaningless. Player 1 cannot gain by deviating and Player 2 deviating would simply lead to Player 1 playing B and so Player 1 should then play A (i.e. they return to the same action pair).

Example

b) The mean payoff vector at such a solution is the Pareto optimal mean payoff vector which gives the greatest sum of mean payoffs while ensuring both of the players at least 2.

Since the set of Pareto optimal solutions is piecewise linear (it is made up of the lines between (2,8) and (6,6) and between (6,6) and (8,2), the maximum sum must come at one of the endpoints of one of these line segments.

The appropriate mean payoff vector is thus (6,6). At such an equilibrium, the players should always play A. If either deviates from this action, then the other should always play B.

Example

It should be noted that this solution cannot be attained using a correlated equilibrium.

It follows that the ability to react to the strategy of other players in a repeated game is a stronger force than communication without the power of making a contract (i.e. the conditions required for attaining a correlated equilibrium), in the sense that the set of mean payoffs possible at an equilibrium of a repeated game contains the set of expected payoffs using a correlated equilibrium in a one-shot game.