

Economics 431

Infinitely repeated games

Let us compare the profit incentives to defect from the cartel in the short run (when the firm is the only defector) versus the long run (when the game is repeated)

The payoffs come from a duopoly forming a cartel on a market with linear demand $p = 30 - Q$, where each firm has a marginal cost $c = 6$. If each firm stays in the cartel, each gets half of monopoly profit, 72. If one firm leaves, the leaving firm produces more than its share of cartel output, and gets 81. If both firms leave, they play a Cournot game, with payoffs of 64 to each player. This is a prisoner's dilemma: whoever leaves first reaps the profit from selling more at high price, but if the other firm leaves, both firms are worse off than they would be in a cartel.

| | | |
|-----|----------------|---------------|
| | S | L |
| S | 72, 72 | 54, <u>81</u> |
| L | <u>81</u> , 54 | 64, <u>64</u> |

If the firms simultaneously decide whether to stay or leave, the unique Nash equilibrium is for both to leave. Suppose that the cartel exists for multiple periods and each period there are two defectors who simultaneously decide whether to stay or leave.

We know from our analysis of the chain store paradox what will happen if the cartel exists for a pre-determined number of periods, say T . If the firms agree to stay in the cartel, they will necessarily both cheat in the last period (it is like your landlord who never returns a security deposit if you move out of town). By backwards induction, we can say that since staying now never affects future payoff. Therefore, every firm will choose an action that maximizes current payoff. The unique subgame perfect equilibrium is to choose L every period.

The problem with finitely repeated games is the so-called end-game problem: no matter how far away the end is, if the end occurs at a certain time, the chain store paradox will work.

However, if we assume that the number of periods is not predetermined, we can avoid end-game problems. If there is some probability that the game continues tomorrow, then actions today *always* have at least *some* ramifications for tomorrow. In contrast, in a finitely repeated game, there is a period with no tomorrow.

We will thus consider infinitely repeated games. They are called so not because it is necessary that the game goes on forever. It is more precise to call them indefinitely repeated games: there is tomorrow with some probability.

As we have seen - a defector has a short-run gain and a long-run loss. How do we compare the payoffs receive in different periods? We use a *discount factor*.

The concept of present value

We compare streams (sequences) of payoffs from different periods using a single number - the present value of a payoff sequence.

- The present value is the sum that we are willing to accept today *instead of* future payoff

Let us compute this sum. What sum would you accept instead of π dollars payable a year from now? Will it be more than π or less than π ? If you have $\frac{\pi}{1+r}$ today, you can put it in the bank for a year and get exactly π a year from now

$$PV = \frac{\pi}{1+r}$$

What if a year from now you are offered a lottery where you win π with probability p and win 0 with probability $1-p$? What sum will you accept today instead of this lottery? You will probably accept something that equal the present value of your expected winning.

$$PV = \frac{p\pi + (1-p)0}{1+r} = \frac{p}{1+r}\pi$$

$$\delta = \frac{p}{1+r}$$

is called discount factor.. It shows you that if the game continues tomorrow with probability p and the interest rate is r , then every dollar received in the tomorrows game will be worth δ dollars payable today.

Example: compute the present value of an infinite sequence of payoffs using the Sum of infinite geometric series

$$1 + \delta + \delta^2 + \delta^3 + \dots = \frac{1}{1-\delta}$$

(10, 10, 10, ...)

$$PV = 10 + 10\delta + 10\delta^2 + \dots = \frac{10}{1-\delta}$$

(0, 2, 1, 1, ...)

$$\begin{aligned} PV &= 0 + 2 \cdot \delta + 1 \cdot \delta^2 + 1 \cdot \delta^3 + \dots = \\ &= 0 + 2\delta + \delta^2 (1 + \delta + \delta^2 + \dots) = 2\delta + \frac{\delta^2}{1-\delta} \end{aligned}$$

(1, 2, 1, 2, 1, 2, ...)

$$\begin{aligned} PV &= 1 + 2\delta + \delta^2 + 2\delta^3 + \delta^4 + 2\delta^5 + \dots = \\ &= (1 + \delta^2 + \delta^4 + \dots) + 2\delta (1 + \delta^2 + \delta^4 + \dots) = \frac{1}{1-\delta^2} + \frac{2\delta}{1-\delta^2} \end{aligned}$$

What is a strategy in an infinitely repeated game?

A strategy should tell you what actions to take in every decision node. But how we describe decision nodes when we have an infinitely large tree?

A *history* of the game is a record of all past actions that the players took. Each history corresponds to a path to a particular decision node (or information set).

Then, when we know the history of the game (all past actions of all players), we know what information set we are in.

For example, suppose that the cartel game

| | | |
|-----|----------------|---------------|
| | S | L |
| S | 72, 72 | 54, <u>81</u> |
| L | <u>81</u> , 54 | 64, <u>64</u> |

is repeated twice. Consider the second period of the game. A history is the record or all past actions taken by players. One example of a history would be S_1L_1 meaning that firm 1 stayed in the first period (S_1) and firm 2 left in the first period. The set of all possible histories H_1 includes all possible action combinations that could have occurred in the first period. This set, not surprisingly, has 4 elements

$$H_1 = \{S_1S_1, S_1L_1, L_1S_1, L_1L_1\}$$

The strategy for the twice repeated cartel game is a rule that tells a player what to do after every history. The strategy must tell a player what to do in the first period and what to do after each of the 4 histories in H_1 . Thus a strategy is a five-letter word, each letter being S or L . For example, $(S_1L_2L_2S_2S_2)$ is a strategy telling a player to play S in the first period (S_1), play L in the second period if the history was S_1S_1 or S_1L_1 and play S in the second period if the history was L_1S_1 or L_1L_1 . Even in this simple game there are $32 = 2^5$ different strategies for either player.

Because of the huge number of strategies, full description of all strategies is impractical if we have a game repeated many times. However, it is not necessary to look at all possible strategies. We can look at strategies that tell the players to take the same action after a whole lot of somehow "similar" histories. A strategy that tells a player to play S in the first period and play S in the second period *if and only if* both players stayed in the first period can be written as $(S_1S_2L_2L_2L_2)$. Notice that the same action L_2 is specified after all histories where some player has left in the past.

What is a subgame perfect equilibrium in an infinitely repeated game?

A subgame perfect strategy tells a player to play the best response to the opponents strategy after every history. That is, *given* the opponent's strategy, a subgame perfect strategy must always (after every history) prescribe an action that is a best response to the opponent's strategy. The action that is a best response must give a player a higher payoff than *any other* action available after this particular history.

Theoretically, we need to go history by history and check every time whether a strategy prescribes a best response. But since the same action may be prescribed after many "similar" histories, many of these checks are redundant.

Example: The trigger strategy in prisoner's dilemma. Let us reconsider the cartel game, which will now be *infinitely* repeated.

Let us divide all histories in two groups:

The first group, call it C , is all the histories that involved both firms playing S in the past. There is only one such history

$$C = \{(SS, SS, SS, \dots)\}$$

The second group, call it P , is all other histories - all histories that involved someone leaving in the past.

Let the firm's strategy be the following:

Start with playing S

After the history where both firms have played only S in the past (i.e. after a history from group C), play S now

After all other histories play L

This strategy describes the "grim trigger" agreement: "I will stay in the cartel today as long as everybody stayed in the past. If not, I will leave and never stay again"

Let us check for which discount factors this strategy is subgame perfect.

We have just two groups of histories to worry about.

Is playing S after C a best response to the opponent's trigger strategy? If it is not, then choosing an alternative action (L) should give more payoff

$$\begin{aligned} \text{Leave now} & : (LS, LL, LL, \dots) \\ \text{Stay now} & : (SS, SS, SS, \dots) \end{aligned}$$

Staying now is better, if payoff from leaving is less.

$$\underbrace{81}_{\text{Get this today}} + \underbrace{\delta \frac{64}{1-\delta}}_{\text{Get 64 forever beginning tomorrow}} < \underbrace{72 \frac{1}{1-\delta}}_{\text{Get 72 forever}}$$

$$\begin{aligned} 81(1-\delta) + 64\delta & < 72 \\ \delta & > \frac{81-72}{81-64} = \frac{\pi_D - \pi_C}{\pi_D - \pi_P} = \frac{9}{17} \end{aligned}$$

After all other histories (group P): the strategy tells to play L (the other player plays L every period after these histories)

$$(LL, LL, LL, \dots), \text{ payoff } \frac{64}{1-\delta}$$

If a player plays S instead, the current payoff is reduced and the future payoff is unaffected

$$(SL, LL, LL, \dots), \text{ payoff } 54 + \delta \frac{64}{1-\delta}$$

Therefore, if $\delta > \frac{9}{17}$, player 1 plays best response after every history. The trigger strategy is subgame perfect. Player 2 is absolutely symmetric. For him, the trigger strategy is also subgame perfect if $\delta > \frac{9}{17}$.

The trigger strategy imposes the harshest possible punishment - it tells the players to stay in the punishment phase forever. Note that both firms using this strategy have an incentive to declare that "bygones are bygones", "forget" the previous history of play and start afresh. In this sense strategies that involve permanent punishments are less satisfactory - why would anyone follow these strategies if both parties have an incentive to re-negotiate the punishment agreement?

An alternative strategy for enforcing cooperation in the repeated game involves punishing the deviating firm for just one period and then forgiving it and going back to cooperation.

The strategy works as follows. All histories of play are subdivided into three groups (or phases):

C - cooperative phase - all histories that involved both players choosing the same action in the every previous period (either LL or SS)

P_1 - punishment phase for player 1 - histories where player 1 has cheated in the previous period (by either leaving the cartel or not accepting the punishment)

P_2 - punishment phase for player 2 - histories where player 2 has cheated in some way in the previous period

Players choose their actions based only on the phase of the game. The candidate equilibrium strategies are:

| Player 1 | | | Player 2 | | |
|----------|--------|-------------------|----------|--------|-------------------|
| Phase | Action | Description | Phase | Action | Description |
| C | S | Cooperate | C | S | Cooperate |
| P_1 | S | Accept punishment | P_1 | L | Impose punishment |
| P_2 | L | Impose punishment | P_2 | S | Accept punishment |

The rule by which the game changes phases depends on the phase it started in and on the actions that players choose in the current period. The table below shows the phase that the game goes into depending on the initial phase and the outcome of the current period

| Initial phase | Current play | | | |
|---------------|--------------|-------|-------|-------|
| | SS | LS | SL | LL |
| C | C | P_1 | P_2 | C |
| P_1 | C | P_1 | C | P_1 |
| P_2 | C | C | P_2 | P_2 |

In words: players cooperate (play S) as long as the game is in the cooperative phase. If player i has cheated, the game changes to the punishment phase P_i - look at the first row of the table above. If the game starts in phase P_i and player i accepts the punishment (i.e. plays S), he is forgiven and the game returns to cooperative phase next period. However, if player i does not accept the punishment, he is not

forgiven, the game stays in the punishment phase next period. That is, the game stays in the punishment phase until the cheater (player i) accepts the punishment, i.e. until player i plays S .

To see if the strategy described is a subgame perfect equilibrium, we must show that there are no profitable deviations after any history, that is, there are no profitable deviation in any phase of the game.

From phase C : The strategy tells player 1 to play S . Then the game will stay in phase C and player 1 will have the payoff of

$$72 + 72\delta + \delta^2 \frac{72}{1 - \delta}.$$

If player 1 deviates to L , he is going to be punished next period and forgiven two periods from now. In other words, if player 1 chooses action L instead of S , the game will go into phase P_1 one period from now and back into phase C to periods from now. The corresponding payoffs are 81 now, 54 next period and 72 in each subsequent period, with the present value of

$$81 + 54\delta + \delta^2 \frac{72}{1 - \delta}.$$

Equilibrium action S has a higher payoff if

$$\begin{aligned} 72 + 72\delta + \delta^2 \frac{72}{1 - \delta} &> 81 + 54\delta + \delta^2 \frac{72}{1 - \delta} \\ 18\delta &> 9, \delta > \frac{1}{2} \end{aligned}$$

From phase P_1 : The strategy tells player 1 to play S , while his opponent will be playing L . Then the game will return to phase C next period and player 1 will have the payoff of

$$54 + 72\delta + \delta^2 \frac{72}{1 - \delta}.$$

If player 1 deviates to L , he is going to be punished again next period and forgiven only two periods from now. That is, he will get 64 now, 54 next period and 72 in each subsequent period, with the present value of

$$64 + 54\delta + \delta^2 \frac{72}{1 - \delta}.$$

In order for player 1 to accept the punishment, doing so must give a higher payoff:

$$\begin{aligned} 54 + 72\delta + \delta^2 \frac{72}{1 - \delta} &> 64 + 54\delta + \delta^2 \frac{72}{1 - \delta} \\ 18\delta &> 10, \delta > \frac{5}{9} > \frac{1}{2} \end{aligned}$$

From phase P_2 . The strategy tells player 1 to impose the punishment by playing L against S . If player 1 changes his action to S , he reduces his current payoff from 81 to 64 and does not affect the future payoff

Note that this strategy works for a smaller range of discount factors than the trigger strategy. The trigger strategy works for $\frac{9}{17} < \delta < 1$. The "punish and forgive" strategy works in a smaller range $\frac{5}{9} < \delta < 1$. This is because temporary punishment is not as harsh as a permanent one of trigger strategy, and therefore it is harder to sustain cooperation with a lesser punishment.