Repeated Games: Infinitely Repeated Games

Threats and promises about future behavior can influence current behavior in repeated relationships.

Reasoning a bit different from finite games

No final state to start backward induction at.

Finite-horizon: if there are multiple Nash equilibria of the stage game G then there may be subgame-perfect outcomes of the repeated game $G(T)$ in which, for any $t < T$, the outcome of stage t is not a Nash equilibrium of G.

Infinitely repeated games: even if the stage game has a unique Nash equilibrium, there may be subgame-perfect outcomes of the infinitely repeated game in which no stage's outcome is a Nash equilibrium of G.

> Begin by studying the infinitely repeated Prisoners' Dilemma.
>
> Then consider the class of infinitely repeated games analogous to the class of finitely repeated games: a static game of complete information, G, is repeated infinitely, with the outcomes of all previous stages observed before the current stage begins

The Prisoners' Dilemma in Figure 2.3.6 is to be repeated infinitely and that, for each t, the outcomes of the $t-1$ preceding plays of the stage game are observed before the t^{th} stage begins.

Summing the payoffs from this infinite sequence of stage games does not provide a useful measure of a player's payoff in the infinitely repeated game.

> Receiving a payoff of 4 in every period is better than receiving a payoff of 1 in every period, but the sum of the payoffs is infinity in both cases.

The discount factor $\delta = 1/(1+r)$ is the value today of a dollar to be received one stage later, where r is the interest rate per stage.

Given a discount factor and a player's payoffs from an infinite sequence of stage games, we can compute the present value of the payoffs-the lump-sum payoff that could be put in the bank now so as to yield the same bank balance at the end of the sequence.

Player 2

	L_2	R_2
L_1	1,1	5,0

R_1	0,5	4,4

Figure 2.3.6.

Definition Given the discount factor δ, the present value of the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \ldots$ is:

$$\pi_1 + \delta\pi_2 + \delta^2\pi_3 + \cdots = \sum_{t=1}^{\infty} \delta^{t-1}\pi_t$$

We can also use δ to reinterpret what we call an infinitely repeated game as a repeated game that ends after a random number of repetitions.

Suppose that after each stage is played a (weighted) coin is flipped to determine whether the game will end.

If the probability is p that the game ends immediately, and therefore $1 - p$ that the game continues for at least one more stage, then a payoff π to be received in the next stage (if it is played) is worth only $(1 - p)\pi/(1 + r)$ before this stage's coin flip occurs.

Likewise, a payoff π to be received two stages from now (if both it and the intervening stage are played) is worth only $(1 - p)^2\pi/(1 + r)^2$ before this stage's coin flip occurs.

Let $\delta = (1 - p)/(1 + r)$.

Then the present value $\pi_1 + \delta\pi_2 + \delta^2\pi_3 + \cdots$ reflects both the time-value of money and the possibility that the game will end.

Consider the infinitely repeated Prisoners' Dilemma in which each player's discount factor is δ

Each player's payoff in the repeated game is the present value of the player's payoffs from the stage games.

We will show that cooperation (R_1, R_2) can occur in every stage of a subgame-perfect outcome of the infinitely repeated game, even though the only Nash equilibrium in the stage game is noncooperation-that is, (L_1, L_2).

The argument is in the spirit of our analysis of the two-stage repeated game with multiple Nash Equilibria

If the players cooperate today then they play a high-payoff equilibrium tomorrow

Otherwise they play a low-payoff equilibrium tomorrow.

The difference is that the high-payoff equilibrium that might be played tomorrow is continued cooperation tomorrow and thereafter (instead of ‘good NE from stage game’).

Player i begins the infinitely repeated game by cooperating and then cooperates in each subsequent stage game if and only if both players have cooperated in every previous stage.

Formally, player $i's$ strategy is:

Play R_i in the first stage. In the t^{th} stage, if the outcome of all $t-1$ preceding stages has been (R_1, R_2) then play R_i; otherwise, play L_i.

This strategy is an example of a trigger strategy, so called because player i cooperates until someone fails to cooperate, which triggers a switch to noncooperation forever after.

If both players adopt this trigger strategy then the outcome of the infinitely repeated game will be (R_1, R_2) in every stage.

If δ is close enough to one then it is a Nash equilibrium of the infinitely repeated game for both players to adopt this strategy.

Such a Nash equilibrium is subgame-perfect, in a sense to be made precise.

Proposition: the trigger strategy is a Nash Equilibrium (in the whole game).

To show that it is a Nash equilibrium of the infinitely repeated game for both players to adopt the trigger strategy…

Assume that player i has adopted the trigger strategy

Then show that, provided δ is close enough to one, it is a best response for player j to adopt the strategy also.

Since player i will play L_i forever once one stage's outcome differs from (R_1, R_2), player j's best response is indeed to play L_j forever once one stage's outcome differs from (R_1, R_2).

It remains to determine player j 's best response in the first stage, and in any stage such that all the preceding outcomes have been (R_1, R_2).

Playing L_j will yield a payoff of 5 this stage but will trigger noncooperation by player i (and therefore also by player j) forever after, so the payoff in every future stage will be 1.

The present value of this sequence of payoffs is

$$5 + \delta \cdot 1 + \delta^2 \cdot 1 + \cdots = 5 + \frac{\delta}{1-\delta}$$

Alternatively, playing R_j will yield a payoff of 4 in this stage and will lead to exactly the same choice between L_j and R_j in the next stage.

Let V denote the present value of the infinite sequence of payoffs player j receives from making this choice optimally (now and every time it arises subsequently).

If playing R_j is optimal, then:

$$V = 4 + \delta V$$

$$V = 4/(1-\delta)$$

Because playing R_j leads to the same choice next stage.

So playing R_j is optimal if and only if

$$\frac{4}{1-\delta} \geq 5 + \frac{\delta}{1-\delta} \tag{2.3.1}$$

$$\delta \geq 1/4$$

In the first stage, and in any stage such that all the preceding outcomes have been (R_1, R_2), player j 's optimal action (given that player i has adopted the trigger strategy) is R_j if and only if $\delta \geq 1/4$.

Combining this observation with the fact that j 's best response is to play L_j forever once one stage's outcome differs from (R_1, R_2), we have that it is a Nash equilibrium for both players to play the trigger strategy if and only if $\delta \geq 1/4$.

Define now a strategy in a repeated game, a subgame in a repeated game, and a subgame-perfect Nash equilibrium in a repeated game.

Stage game $G = \{A_1, \ldots, A_n; u_1, \ldots, u_n\}$:

> Static game of complete information in which players 1 through n simultaneously choose actions a_1 through a_n from the action spaces A_1 through A_n, respectively, and payoffs are $u_1(a_1, \ldots, a_n)$ through $u_n(a_1, \ldots, a_n)$.

We had before a finitely repeated game $G(T)$ based on G.

Define the analogous infinitely repeated game.

Definition: infinitely repeated game

$G(\infty, \delta)$ denotes the infinitely repeated game in which G is repeated forever and the players share the discount factor δ.

For each t, the outcomes of the $t-1$ preceding plays of the stage game are observed before the t^{th} stage begins.

Each player's payoff in $G(\infty, \delta)$ is the present value of the player's payoffs from the infinite sequence of stage games.

In any game (repeated or otherwise), a player's strategy is a complete plan of action:

It specifies a feasible action for the player in every contingency in which the player might be called upon to act.

In a dynamic game, however, a strategy is more complicated than a simple action.

In the finitely repeated game $G(T)$ or the infinitely repeated game $G(\infty, \delta)$, the history of play through stage t is the record of the players' choices in stages 1 through t.

The players might have chosen $(a_{11}, \ldots, a_{n1})$ in stage 1, $(a_{12}, \ldots, a_{n2})$ in stage 2, ..., and $(a_{1t}, \ldots, a_{nt})$ in stage t, for example, where for each player i and stage s the action a_{is} belongs to the action space A_i.

Definition In the finitely repeated game $G(T)$ or the infinitely repeated game $G(\infty, \delta)$, a player's strategy specifies the action the player will take in each stage, for each possible history of play through the previous stage.

Subgames

Definition In the finitely repeated game $G(T)$, a subgame beginning at stage $t + 1$ is the repeated game in which G is played $T - t$ times, denoted $G(T - t)$.

There are many subgames that begin at stage $t + 1$, one for each of the possible histories of play through stage t.

In the infinitely repeated game $G(\infty, \delta)$, each subgame beginning at stage $t + 1$ is identical to the original game $G(\infty, \delta)$.

As in the finite-horizon case, there are as many subgames beginning at stage $t + 1$ of $G(\infty, \delta)$ as there are possible histories of play through stage t.

Note that the t^{th} stage of a repeated game taken on its own is not a subgame of the repeated game (assuming $t < T$ in the finite case).

Proposition: The trigger strategy in the PD is a subgame perfect Nash equilibrium.

We must show that the trigger strategies constitute a Nash equilibrium on every subgame of that infinitely repeated game.

Recall that every subgame of an infinitely repeated game is identical to the game as a whole.

In the trigger-strategy Nash equilibrium of the infinitely repeated Prisoners' Dilemma, these subgames can be grouped into two classes:

(i) Subgames in which all the outcomes of earlier stages have been (R_1, R_2)

(ii) Subgames in which the outcome of at least one earlier stage differs from (R_1, R_2).

If the players adopt the trigger strategy for the game as a whole, then

(i) the players' strategies in a subgame in the first class are again the trigger strategy, which we have shown to be a Nash equilibrium of the game as a whole

(ii) and the players' strategies in a subgame in the second class are simply to repeat the stage-game equilibrium (L_1, L_2) forever, which is also a Nash equilibrium of the game as a whole.

Thus, the trigger-strategy Nash equilibrium of the infinitely repeated Prisoners' Dilemma is subgame-perfect.

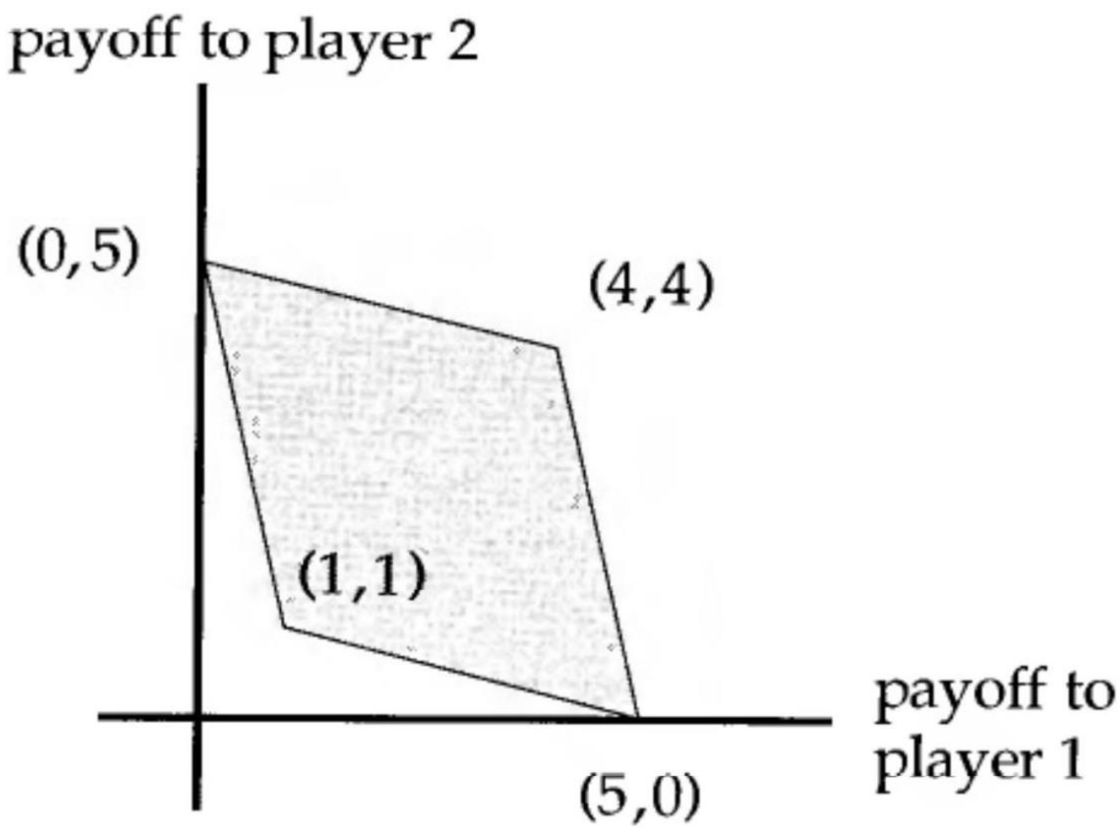

Figure 2.3.7.

We next apply analogous arguments in the infinitely repeated game $G(\infty, \delta)$.

These arguments lead to Friedman's (1971) Theorem for infinitely repeated games.

To state the theorem, we need two final definitions.

First, we call the payoffs $(x_1, \ldots, x_n)$ feasible in the stage game G if they are a convex combination (i.e., a weighted average, where the weights are all nonnegative and sum to one) of the pure-strategy payoffs of G.

> The set of feasible payoffs for the Prisoners' Dilemma in Figure 2.3.6 is the shaded region in Figure 2.3.7.
>
> The pure-strategy payoffs $(1,1)$, $(0,5)$, $(4,4)$, and $(5,0)$ are feasible.
>
> Other feasible payoffs include the pairs (x, x) for $1 < x < 4$, which result from weighted averages of $(1,1)$ and $(4,4)$, and the pairs (y, z) for $y + z = 5$ and $0 < y < 5$, which result from weighted averages of $(0,5)$ and $(5,0)$.
>
> The other pairs in (the interior of) the shaded region in Figure 2.3.7 are weighted averages of more than two pure-strategy payoffs.
>
> To achieve a weighted average of pure-strategy payoffs, the players could use a public randomizing device: by playing (L_1, R_2) or (R_1, L_2) depending on a flip of a (fair) coin, for example, they achieve the expected payoffs $(2.5, 2.5)$.

Second: we need a rescaling of the players' payoffs.

> We continue to define each player's payoff in the infinitely repeated game $G(\infty, \delta)$ to be the present value of the player's infinite sequence of stagegame payoffs
>
> But it is more convenient to express this present value in terms of the average payoff from the same infinite sequence of stage-game payoffs-the payoff that would have to be received in every stage so as to yield the same present value.
>
> Let the discount factor be δ.
>
> Suppose the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \ldots$ has a present value of V.
>
> If the payoff π were received in every stage, the present value would be $\pi/(1-\delta)$.
>
> For π to be the average payoff from the infinite sequence $\pi_1, \pi_2, \pi_3 \ldots$ with discount factor δ, these two present values must be equal, so $\pi = V(1-\delta)$.
>
> That is, the average payoff is $(1-\delta)$ times the present value.

Definition Given the discount factor δ, the average payoff of the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \ldots$ is

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\pi_t$$

The advantage of the average payoff over the present value is that the former is directly comparable to the payoffs from the stage game.

In the Prisoners' Dilemma in Figure 2.3.6, both players might receive a payoff of 4 in every period.

> Such an infinite sequence of payoffs has an average payoff of 4 but a present value of $4/(1-\delta)$.
>
> Since the average payoff is just a rescaling of the present value, however, maximizing the average payoff is equivalent to maximizing the present value.

We are at last ready to state the main result in our discussion of infinitely repeated games:

Theorem (Friedman 1971)

Let G be a finite, static game of complete information.

Let $(e_1, \dots, e_n)$ denote the payoffs from a Nash equilibrium of G

Let $(x_1, \dots, x_n)$ denote any other feasible payoffs from G.

If $x_i > e_i$ for every player i and if δ is sufficiently close to one, then there exists a subgame-perfect Nash equilibrium of the infinitely repeated game $G(\infty, \delta)$ that achieves $(x_1, \dots, x_n)$ as the average payoff.

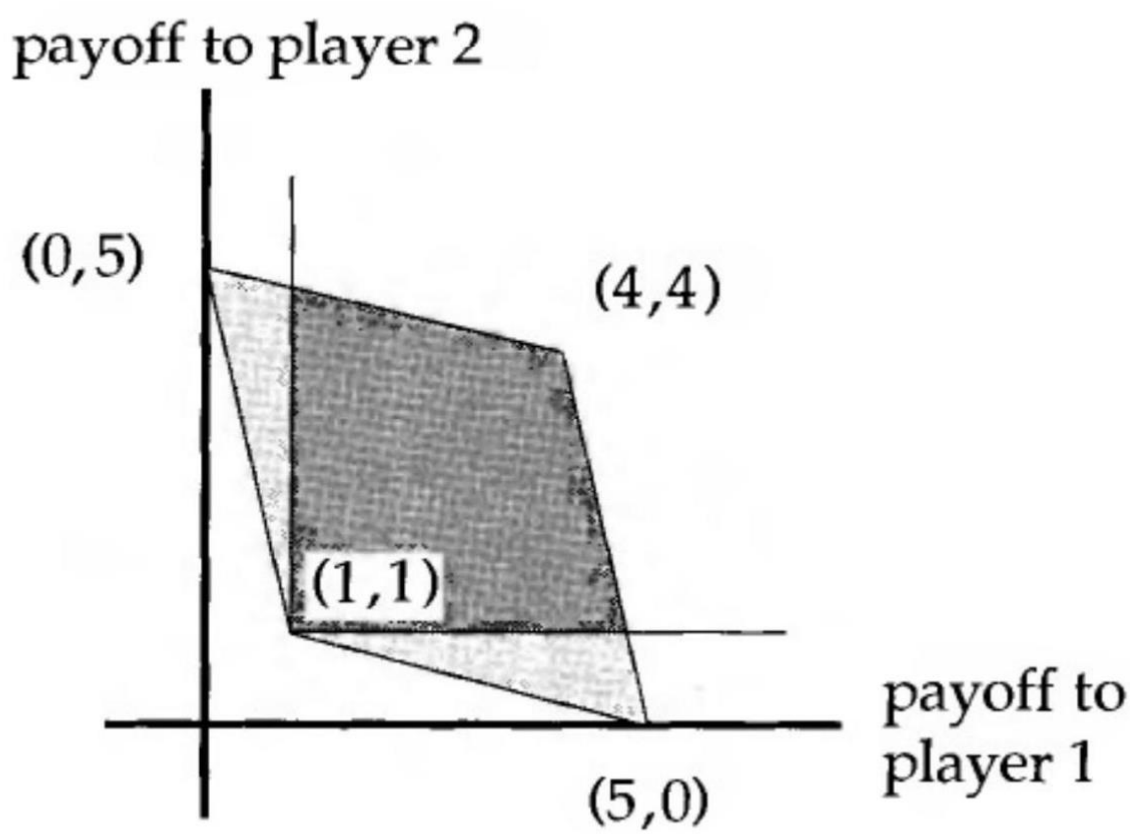

Figure 2.3.8

In the context of the Prisoners' Dilemma in Figure 2.3.6, Friedman's Theorem guarantees that any point in the cross-hatched region in Figure 2.3.8 can be achieved as the average payoff in a subgame-perfect Nash equilibrium of the repeated game, provided the discount factor is sufficiently close to one.

One remark

In the (one-shot) Prisoners' Dilemma in Figure 2.3.6, player i can guarantee receiving at least the Nash equilibrium payoff of 1 by playing L_i.

In a one-shot Cournot duopoly game, in contrast, a firm cannot guarantee receiving the Nash-equilibrium profit by producing the Nash-equilibrium quantity

> The only profit a firm can guarantee receiving is zero, by producing zero

Given an arbitrary stage game G, let r_i denote player i 's reservation payoff-the largest payoff player i can guarantee receiving, no matter what the other players do.

Hence $r_i \leq e_i$ (where e_i is player i 's Nash equilibrium payoff used in Friedman's Theorem)

> If r_i were greater than e_i, it would not be a best response for player i to play his or her Nash-equilibrium strategy.

In the Prisoners' Dilemma, $r_i = e_i$, but in the Cournot duopoly game (and typically), $r_i < e_i$.

Fudenberg and Maskin (1986) show that for two-player games, the reservation payoffs (r_1, r_2) can replace the equilibrium payoffs (e_1, e_2) in the statement of Friedman's Theorem.

That is, if (x_1, x_2) is a feasible payoff from G, with $x_i > r_i$ for each i, then for δ sufficiently close to one there exists a subgame-perfect Nash equilibrium of $G(\infty, \delta)$ that achieves (x_1, x_2) as the average payoff, even if $x_i < e_i$ for one or both of the players.

For games with more than two players, Fudenberg and Maskin provide a mild condition under which the reservation payoffs $(r_1, \ldots, r_n)$ can replace the equilibrium payoffs $(e_1, \ldots, e_n)$ in the statement of the Theorem.

Another Remark

What average payoffs can be achieved by subgame-perfect Nash equilibria when the discount factor is not "sufficiently close to one"?

One way to approach this question is to consider a fixed value of δ and determine the average payoffs that can be achieved if the players use trigger strategies that switch forever to the stage-game Nash equilibrium after any deviation.

Smaller values of δ make a punishment that will begin next period less effective in deterring a deviation this period.

Nonetheless, the players typically can do better than simply repeating a stage-game Nash equilibrium.

A second approach, pioneered by Abreu (1988), is based on the idea that the most effective way to deter a player from deviating from a proposed strategy is to threaten to administer the strongest credible punishment should the player deviate

> I.e., threaten to respond to a deviation by playing the subgame-perfect Nash equilibrium of the infinitely repeated game that yields the lowest payoff of all such equilibria for the player who deviated.

In most games, switching forever to the stage-game Nash equilibrium is not the strongest credible punishment

> Hence some average payoffs can be achieved using Abreu's approach that cannot be achieved using the trigger-strategy approach.

In the Prisoners' Dilemma, however, the stage-game Nash equilibrium yields the reservation payoffs (that is, $e_i = r_i$), so the two approaches are equivalent.

Proof of Friedman's Theorem

Let $(a_{e1}, \dots, a_{en})$ be the Nash equilibrium of G that yields the equilibrium payoffs $(e_1, \dots, e_n)$.

Let $(a_{x1}, \dots, a_{xn})$ be the collection of actions that yields the feasible payoffs $(x_1, \dots, x_n)$.

> (The latter notation is only suggestive because it ignores the public randomizing device typically necessary to achieve arbitrary feasible payoffs.)

Consider the following trigger strategy for player i :

> Play a_{xi} in the first stage.
>
> In the t^{th} stage, if the outcome of all $t-1$ preceding stages has been $(a_{x1}, \dots, a_{xn})$ then play a_{xi}
>
> Otherwise, play a_{ei}.

If both players adopt this trigger strategy then the outcome of every stage of the infinitely repeated game will be ($a_{x1}, \dots, a_{xn}$), with (expected) payoffs $(x_1, \dots, x_n)$.

We first argue that if δ is close enough to one, then it is a Nash equilibrium of the repeated game for the players to adopt this strategy.

We then argue that such a Nash equilibrium is subgame-perfect.

Suppose that all the players other than player i have adopted this trigger strategy.

Since the others will play $(a_{e1}, \ldots, a_{e,i-1}, a_{e,i+1}, \ldots, a_{en})$ forever once one stage's outcome differs from $(a_{x1}, \ldots, a_{xn})$, player i 's best response is to play a_{ei} forever once one stage's outcome differs from $(a_{x1}, \ldots, a_{xn})$.

It remains to determine player i 's best response in the first stage, and in any stage such that all the preceding outcomes have been $(a_{x1}, \ldots, a_{xn})$.

Let a_{di} be player i' s best deviation from $(a_{x1}, \ldots, a_{xn})$.

That is, a_{di} solves

$$\max_{a_i \in A_i} u_i(a_{x1}, \ldots, a_{x,i-1}, a_i, a_{x,i+1}, \ldots, a_{xn})$$

Let d_i be i 's payoff from this deviation:

$$d_i = u_i(a_{x1}, \ldots, a_{x,i-1}, a_{di}, a_{x,i+1}, \ldots, a_{xn})$$

(Again, we ignore the role of the randomizing device: the best deviation and its payoff may depend on which pure strategies the randomizing device has prescribed.)

We have:

$$d_i \geq x_i = u_i(a_{x1}, \ldots, a_{x,i-1}, a_{xi}, a_{x,i+1}, \ldots, a_{xn}) > e_i = u_i(a_{e1}, \ldots, a_{en}).$$

Playing a_{di} will yield a payoff of d_i at this stage but will trigger $(a_{e1}, \ldots, a_{e,i-1}, a_{e,i+1}, \ldots, a_{en})$ by the other players forever after, to which the best response is a_{ei} by player i, so the payoff in every future stage will be e_i.

The present value of this sequence of payoffs is

$$d_i + \delta \cdot e_i + \delta^2 \cdot e_i + \cdots = d_i + \frac{\delta}{1-\delta} e_i$$

(Since any deviation triggers the same response by the other players, the only deviation we need to consider is the most profitable one.)

Alternatively, playing a_{xi} will yield a payoff of x_i this stage and will lead to exactly the same choice between a_{di} and a_{xi} in the next stage.

Let V_i denote the present value of the stage-game payoffs player i receives from making this choice optimally (now and every time it arises subsequently).

If playing a_{xi} is optimal, then

$$V_i = x_i + \delta V_i$$

or $V_i = x_i/(1-\delta)$

If playing a_{di} is optimal, then

$$V_i = d_i + \frac{\delta}{1-\delta} e_i$$

as derived previously.

(Assume that the randomizing device is serially uncorrelated. It then suffices to let d_i be the highest of the payoffs to player i 's best deviations from the various pure-strategy combinations prescribed by the randomizing device.)

So playing a_{xi} is optimal if and only if

$$\frac{x_i}{1-\delta} \geq d_i + \frac{\delta}{1-\delta} e_i$$

or

$$\delta \geq \frac{d_i - x_i}{d_i - e_i}$$

Thus, in the first stage, and in any stage such that all the preceding outcomes have been $(a_{x1}, \dots, a_{xn})$, player i 's optimal action (given that the other players have adopted the trigger strategy) is a_{xi} if and only if:

$$\delta \geq (d_i - x_i)/(d_i - e_i).$$

Combining this observation with the fact that i 's best response is to play a_{ei} forever once one stage's outcome differs from $(a_{x1}, \dots, a_{xn})$, we have that it is a Nash equilibrium for all the players to play the trigger strategy if and only if

$$\delta \geq \max_i \frac{d_i - x_i}{d_i - e_i}$$

Since $d_i \geq x_i > e_i$, it must be that $(d_i - x_i)/(d_i - e_i) < 1$ for every i, so the maximum of this fraction across all the players is also strictly less than one.

It remains to show that this Nash equilibrium is subgame perfect.

That is, the trigger strategies must constitute a Nash equilibrium in every subgame of $G(\infty, \delta)$.

Recall that every subgame of $G(\infty, \delta)$ is identical to $G(\infty, \delta)$ itself.

In the trigger-strategy Nash equilibrium, these subgames can be grouped into two classes:

(i) subgames in which all the outcomes of earlier stages have been $(a_{x1}, \dots, a_{xn})$
(ii) subgames in which the outcome of at least one earlier stage differs from $(a_{x1}, \dots, a_{xn})$.

If the players adopt the trigger strategy for the game as a whole, then

(i) the players' strategies in a subgame in the first class are again the trigger strategy, which we have just shown to be a Nash equilibrium of the game as a whole

(ii) the players' strategies in a subgame in the second class are simply to repeat the stage-game equilibrium $(a_{e1}, \dots, a_{en})$ forever, which is also a Nash equilibrium of the game as a whole.

Thus, the trigger-strategy Nash equilibrium of the infinitely repeated game is subgame-perfect.

Collusion between Cournot Duopolists

Friedman (1971) was the first to show that cooperation could be achieved in an infinitely repeated game by using trigger strategies that switch forever to the stage-game Nash equilibrium following any deviation.

The original application was to collusion in a Cournot oligopoly, as follows.

Static Cournot game:

Aggregate quantity on the market: $Q = q_1 + q_2$

Market-clearing price is $P(Q) = a - Q$, assuming $Q < a$.

Each firm has a marginal cost of c and no fixed costs.

Firms choose quantities simultaneously.

In the unique Nash equilibrium, each firm produces:

$$q_C = (a - c)/3$$

Equilibrium aggregate quantity $2(a - c)/3$ exceeds the monopoly quantity, $q_m \equiv (a - c)/2$,

Hence both firms would be better off if each produced half the monopoly quantity, $q_i = q_m/2$.

Consider the infinitely repeated game based on this Cournot stage game when both firms have the discount factor δ.

Trigger strategy:

Produce half the monopoly quantity, $q_m/2$, in the first period. In the t^{th} period, produce $q_m/2$ if both firms have produced $q_m/2$ in each of the $t - 1$ previous periods; otherwise, produce the Cournot quantity, q_C.

The profit to one firm when both produce $q_m/2$ is $\frac{\pi_m}{2} = (a - c)^2/8$, which we will denote by $\pi_m/2$.

The profit to one firm when both produce q_C is $\pi_C = (a - c)^2/9$

If firm i is going to produce $q_m/2$ this period then the quantity that maximizes firm j 's profit this period solves

$$\max_{q_j} \left(a - q_j - \frac{1}{2} q_m - c\right) q_j$$

The solution is $q_j = 3(a-c)/8$, with associated profit of $\pi_d = 9(a-c)^2/64$ (" d " for deviation).

Thus, it is a Nash equilibrium for both firms to play the trigger strategy given earlier provided that

$$\frac{1}{1-\delta} \cdot \frac{1}{2} \pi_m \geq \pi_d + \frac{\delta}{1-\delta} \cdot \pi_C \tag{2.3.2}$$

Substituting the values of π_m, π_d, and π_C into (2.3.2) yields $\delta \geq 9/17$.

For the same reasons as in the previous section, this Nash equilibrium is subgame-perfect.

We can also ask what the firms can achieve if $\delta < 9/17$.

First approach:

We first determine, for a given value of δ, the most-profitable quantity the firms can produce if they both play trigger strategies that switch forever to the Cournot quantity after any deviation.

We know that such trigger strategies cannot support a quantity as low as half the monopoly quantity

> But for any value of δ it is a subgame-perfect Nash equilibrium simply to repeat the Cournot quantity forever.
>
> Therefore, the most-profitable quantity that trigger strategies can support is between $q_m/2$ and q_C.

To compute this quantity, consider the following trigger strategy:

> Produce q^* in the first period.
>
> In the t^{th} period, produce q^* if both firms have produced q^* in each of the $t-1$ previous periods
>
> Otherwise, produce the Cournot quantity, q_C.

The profit to one firm if both play q^* is $\pi^* = (a - 2q^* - c)q^*$.

If firm i is going to produce q^* this period, then the quantity that maximizes firm j 's profit this period solves

$$\max_{q_j} (a - q_j - q^* - c) q_j$$

Solution is $q_j = (a - q^* - c)/2$

Associated profit of $\pi_d = (a - q^* - c)^2/4$

It is a Nash equilibrium for both firms to play the trigger strategy given above if and only if:

$$\frac{1}{1-\delta} \cdot \pi^* \geq \pi_d + \frac{\delta}{1-\delta} \cdot \pi_C$$

Solving the resulting quadratic in q^* shows that the lowest value of q^* for which the trigger strategies given above are a subgameperfect Nash equilibrium is

$$q^* = \frac{9 - 5\delta}{3(9 - \delta)}(a - c)$$

which is monotonically decreasing in δ, approaching $q_m/2$ as δ approaches $9/17$ and approaching q_C as δ approaches zero.

Second approach (strongest credible punishment – Abreu 1986):

We show that Abreu's approach can achieve the monopoly outcome in our model when $\delta = 1/2$ (which is less than 9/17).

Abreu (1986) applies this idea to Cournot models more general than ours using an arbitrary discount factor

Consider the following "two phase" (or "carrot-and-stick") strategy:

Produce half the monopoly quantity, $q_m/2$, in the first period.

In the t^{th} period, produce $q_m/2$ if both firms produced $q_m/2$ in period $t-1$, produce $q_m/2$ if both firms produced x in period $t-1$

Otherwise produce x.

This strategy involves a (one-period) punishment phase in which the firm produces x and a (potentially infinite) collusive phase in which the firm produces $q_m/2$.

If either firm deviates from the collusive phase, then the punishment phase begins.

If either firm deviates from the punishment phase, then the punishment phase begins again.

If neither firm deviates from the punishment phase, then the collusive phase begins again.

The profit to one firm if both produce x is $(a - 2x - c)x$, which we will denote by $\pi(x)$.

Let $V(x)$ denote the present value of receiving $\pi(x)$ this period and half the monopoly profit forever after:

$$V(x) = \pi(x) + \frac{\delta}{1-\delta} \cdot \frac{1}{2}\pi_m$$

If firm i is going to produce x this period, then the quantity that maximizes firm j 's profit this period solves

$$\max_{q_j}(a - q_j - x - c)q_j$$

Solution: $q_j = (a - x - c)/2$

Associated profit: $\pi_{dp}(x) = (a - x - c)^2/4$ (dp stands for deviation from the punishment).

If both firms play the two-phase strategy above, then the subgames in the infinitely repeated game can be grouped into two classes:

(i) collusive subgames, in which the outcome of the previous period was either $(q_m/2, q_m/2)$ or (x, x)
(ii) punishment subgames, in which the outcome of the previous period was neither $(q_m/2, q_m/2)$, nor (x, x).

For it to be a subgame-perfect Nash equilibrium for both firms to play the two-phase strategy, it must be a Nash equilibrium to obey the strategy in each class of subgames.

In the collusive subgames, each firm must prefer to receive half the monopoly profit forever than to receive π_d this period and the punishment present value $V(x)$ next period:

$$\frac{1}{1-\delta} \cdot \frac{1}{2}\pi_m \geq \pi_d + \delta V(x) \tag{2.3.3}$$

In the punishment subgames, each firm must prefer to administer the punishment than to receive π_{dp} this period and begin the punishment again next period:

$$V(x) \geq \pi_{dp}(x) + \delta V(x) \tag{2.3.4}$$

Substituting for $V(x)$ in (2.3.3) yields

$$\delta\left(\frac{1}{2}\pi_m - \pi(x)\right) \geq \pi_d - \frac{1}{2}\pi_m$$

That is, the gain this period from deviating must not exceed the discounted value of the loss next period from the punishment.

(Provided neither firm deviates from the punishment phase, there is no loss after next period, since the punishment ends and the firms return to the monopoly outcome, as though there had been no deviation.)

Likewise, (2.3.4) can be rewritten as

$$\delta\left(\frac{1}{2}\pi_m - \pi(x)\right) \geq \pi_{dp} - \pi(x)$$

with an analogous interpretation.

For $\delta = 1/2$, (2.3.3) is satisfied provided $x/(a-c)$ is not between $1/8$ and $3/8$, and (2.3.4) is satisfied if $x/(a-c)$ is between $3/10$ and $1/2$.

Thus, for $\delta = 1/2$, the two-phase strategy achieves the monopoly outcome as a subgameperfect Nash equilibrium provided that $3/8 \leq x/(a-c) \leq 1/2$.

Other Dynamic Oligopoly Models

There are many other models of dynamic oligopoly that enrich the simple model developed here.

Two classes of such models:

State-variable models

Imperfect-monitoring models.

Rotemberg and Saloner (1986, and Problem 2.14) study collusion over the business cycle by allowing the intercept of the demand function to fluctuate randomly across periods.

In each period, all firms observe that period's demand intercept before taking their actions for that period;

In other applications, the players could observe the realization of another state variable at the beginning of each period.

The incentive to deviate from a given strategy thus depends both on the value of demand this period and on the likely realizations of demand in future periods.

(Rotemberg and Saloner assume that demand is independent across periods, so the latter consideration is independent of the current value of demand, but later authors have relaxed this assumption.)

Green and Porter (1984) study collusion when deviations cannot be detected perfectly:

Rather than observing the other firms' quantity choices, each firm observes only the market-clearing price, which is buffeted by an unobservable shock each period.

In this setting, firms cannot tell whether a low market-clearing price occurred because one or more firms deviated or because there was an adverse shock.

Green and Porter examine trigger-price equilibria, in which any price below a critical level triggers a punishment period during which all firms play their Cournot quantities.

In equilibrium, no firm ever deviates.

Nonetheless, an especially bad shock can cause the price to fall below the critical level, triggering a punishment period.

Since punishments happen by accident, infinite punishments of the kind considered in the trigger-strategy analysis in this section are not optimal.

Two-phase strategies of the kind analyzed by Abreu might seem promising

Abreu, Pearce, and Stacchetti (1986) show that they can be optimal.