22 REPEATED GAMES AND REPUTATION

People often interact in ongoing relationships. For example, most employment relationships last a long time. Countries competing over tariff levels know that they will be affected by one another's policies far into the future. Firms in an industry recognize that they are not playing a static game but one in which they compete every day over time. In all of these dynamic situations, the way in which a party behaves at any given time is influenced by what this party and others did in the past. In other words, players *condition* their decisions on the history of their relationship. An employee may choose to work diligently only if his employer gave him a good bonus in the preceding month. One country may set a low import tariff only if its trading partners had maintained low tariffs in the past. A firm may elect to match its competitor's price by setting its price each day equal to the competitor's price of the preceding day.

When deciding how to behave in an ongoing relationship, one must consider how one's behavior will influence the actions of others in the future. Suppose I am one of your employees and that our history is one of cooperation. Since you hired me, I have been a loyal and hard-working employee, and you have given me a generous bonus each month (above my salary). Today I am considering whether to work hard or, alternatively, to neglect my duties in favor of playing video games on the office computer. For me, shirking has an immediate reward—I get to avoid expending effort on the job. But you will soon learn of my indolence, either through your monitoring activities or through an observed decrease in my productivity. *Your future behavior* (in particular, whether to give me bonuses each month) may very well be influenced by *what I do today*.

For instance, after observing that I have shirked, you might choose to discontinue my monthly bonus. You might say to yourself, "By misbehaving, Watson has lost my trust; I doubt that he will work diligently ever again and therefore I will pay him no more bonuses." Anticipating such a response, I may decide that spending the workday playing chess over the Internet against someone in New Zealand is not such a good idea. Neglecting my duties may yield an immediate gain (relaxation today), but it leads to a greater loss in the future (no more bonuses each month). As this story suggests, people sometimes have an incentive to forego small immediate gains because of the threat of future retaliation by others.

The term "reputation" is often used to describe how a person's past actions affect future beliefs and behavior. If I have always worked diligently on the job, people would say that I have "established a reputation for being a hard worker." If I shirk today, then tomorrow people would say that I have "destroyed my good reputation." Often, those who nurture good reputations are trusted and rewarded; people with bad reputations are punished. As the employment story indicates, the concern for reputation may motivate parties to cooperate with one another, even if such behavior requires foregoing short-term gains. One of the great achievements of game theory is that it provides a framework for understanding how such a reputation mechanism can support cooperation.

The best way to study the interaction between immediate gains and long-term incentives is to examine a *repeated game*. A repeated game is played over discrete periods of time (period 1, period 2, and so on). We let t denote any given period and let T denote the total number of periods in the repeated game. T can be a finite number or it can be infinity, which means the players interact perpetually over time. In each period, the players play a static *stage game*, whereby they simultaneously and independently select actions. These actions lead to a stage-game payoff for the players. The stage game can be denoted by $\{A, u\}$, where

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

is the set of action profiles and $u_i(a)$ is player i's stage-game payoff when profile a is played. The same stage game is played in each period. Furthermore, we assume that in each period t, the players have observed the *history* of play—that is, the sequence of action profiles—from the first period through period $t - 1$. The payoff of the entire game is defined as the sum of the stage-game payoffs in periods 1 through T. We sometimes assume that players discount the future, in which case we include a discount factor in the payoff specification (recall the analysis of discounting in Chapter 19).

A TWO-PERIOD REPEATED GAME

Suppose players 1 and 2 interact over two periods, called periods 1 and 2 (so $T = 2$). In each period, they play the stage game depicted in Figure 22.1. Assume that the payoff for the entire game is the sum of the stage-game payoffs in the two periods. For instance, if (A, X) is played in the first period and (B, Y) is played in the second period, then player 1's payoff is $4 + 2 = 6$ and player

FIGURE 22.1
Stage game, repeated once ($T = 2$).

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

2's payoff is $3 + 1 = 4$. Figure 22.2(a) graphs the set of possible repeated game payoffs. Every point on the graph corresponds to the sum of two stage-game payoff vectors. For example, the payoff vector (3, 5) can be attained if (A, Z) is played in the first period and (B, Y) is played in the second period; the same payoff results if (B, Y) is played in the first period, followed by (A, Z).

This two-period repeated game has a large extensive-form representation, so I have not drawn it here. The extensive form starts with simultaneous selection of actions in the first period: player 1 chooses between A and B, and player 2 selects X, Y, or Z. Then, having observed each other's first-period actions, the players again select from {A, B} and {X, Y, Z}. Because each player knows what happened in the first period, his choice in the second period can be conditioned on this information. For example, player 1 may decide to pick A in the second period if and only if (A, X) or (B, Y) was played in the first period; otherwise, he picks B. As usual, the players' information is represented by information sets. Because there are six possible outcomes of first-period interaction, each player

FIGURE 22.2
Feasible repeated-game payoffs.

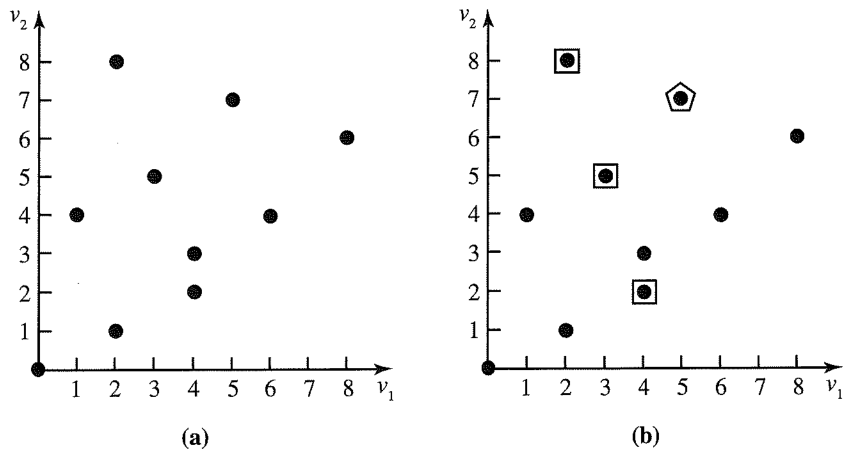


FIGURE 22.3
The subgame following (A, Z).

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

has six information sets in the second period. In other words, each player has *six* decisions to make in period 2: what to do if the outcome of period 1 was (A, X), what to do if the outcome of period 1 was (A, Y), and so forth.

This game has a large set of strategies, making the analysis of rationality a bit daunting. However, the analysis is quite illuminating, so read on. Let us look for pure-strategy, subgame perfect Nash equilibria. We can simplify the search by recognizing one important thing: every equilibrium must specify that in the *second* period, the players select an action profile that is a Nash equilibrium of the stage game.

To see why this is so, recall that subgame perfection requires equilibrium in *every subgame*. In the repeated game at hand, a different subgame is initiated following every different action profile in period 1. For example, consider what happens in the event that the players choose (A, Z) in the first period. Then, knowing that (A, Z) was the outcome of first-period interaction, the players proceed to a subgame in which they simultaneously select actions again. Their total payoff will be (1, 4) plus whatever the payoff vector is in the second play of the stage game. Thus, following the play of (A, Z) in the first period, the subgame is described by the matrix in Figure 22.3. I constructed this matrix by adding the payoff vector (1, 4) to each of the cells in the stage game (compare Figures 22.1 and 22.3). Thus, Figure 22.3 describes the possible *continuation payoffs* following play of (A, Z) in the first period. You should verify that the subgame has two Nash equilibria, (B, Y) and (A, Z). Therefore, a subgame perfect equilibrium must specify that either (B, Y) or (A, Z) be played in the second period if (A, Z) is played in the first period.

We can say more. Because the subgame matrix is formed by adding the same payoff vector to each cell of the stage-game matrix, the players' preferences over action profiles in the subgame *are exactly the same* as their preferences in an isolated play of the stage game. In other words, writing the continuation payoffs net of the stage-game payoff of period 1, the subgame is equivalent to the stage game, and thus they have exactly the same Nash equilibria.[1] In fact, every

[1]You should verify that (A, Z) and (B, Y) are the Nash equilibria of the stage game.

subgame starting in period 2 has the same set of Nash equilibria because these games have matrices like that in Figure 22.3—where the stage-game payoff in the first period is added to every cell of the stage game. Another way of thinking about this is that once the first period is over, the payoffs from the first period are *sunk*. Whatever the players obtain in the second period, it is in addition to what they have already received in the first. Thus a subgame perfect equilibrium must specify that the players select a Nash equilibrium in the stage game in period 2, whatever happens in period 1. To save ink, I use the phrase "stage Nash profile" to refer to a Nash equilibrium in the stage game.

Knowing that a stage Nash profile will be played in the second period, we can turn our attention to two other matters: (1) action choices in the first period, and (2) how behavior in the first period determines *which* of the two stage Nash equilibria will be played in the second period.

First consider subgame perfect equilibria that specify that a stage Nash profile be played in the first period (as well as in the second). Here is one such strategy profile: The players are instructed to choose action profile (A, Z) in the first period and then, regardless of the outcome of the first period, they are to choose (A, Z) in the second period. Because each player has six information sets in the second period (six potential decisions to make), the phrase "regardless of the outcome of the first period" is crucial; it means that, even if one or both of the players deviate from (A, Z) in the first period, they are supposed to play (A, Z) in the second. You can easily verify that this strategy profile is a subgame perfect equilibrium—neither player can gain by deviating in either or both periods, given the other player's strategy. In this equilibrium, player 1 obtains $1 + 1 = 2$ and player 2 gets $4 + 4 = 8$. This payoff vector is one of those boxed in Figure 22.2(b).

Any combination of stage Nash profiles can be supported as a subgame perfect equilibrium outcome. For example, "choose (A, Z) in the first period and then, regardless of the first-period outcome, choose (B, Y) in the second period" is a subgame perfect equilibrium; it yields the payoff vector (3, 5). The payoffs of equilibria that specify stage Nash profiles in both periods are all boxed in Figure 22.2(b). I recommend reviewing the various combinations of stage Nash profiles and verifying that the associated equilibrium payoffs are boxed in Figure 22.2(b). As the example intimates, the following general result holds:

> **Result:** Consider any repeated game. Any sequence of stage Nash profiles can be supported as the outcome of a subgame perfect Nash equilibrium.

It probably does not surprise you that stage Nash profiles can be supported as equilibrium play. A more interesting question is whether there are equilibria stipulating actions that are *not* stage Nash profiles. In fact, the answer is "yes," as the two-period example at hand illustrates. Consider the following strategy profile:

> Select (A, X) in the first period and then, as long as player 2 did not deviate from X, select (A, Z) in the second period; if player 2 deviated by playing Y or Z in the first period, then play (B, Y) in the second period.

This strategy profile prescribes that the players' second-period actions depend on what player 2 did in period 1. By playing X in the first period, player 2 establishes a reputation for cooperating; in this case, he is rewarded in the second period as the players coordinate on the stage Nash profile that is more favorable to him. In contrast, if player 2 deviates by, say, choosing Z in the first period, then he is branded a "cheater." In this case, his punishment is that the players coordinate on (B, Y) in the second period.

To verify that this strategy profile is a subgame perfect equilibrium, we must check each player's incentives. Suppose player 1 behaves as prescribed and consider the incentives of player 2. If player 2 goes along with the strategy prescription, he obtains 3 in the first period and 4 in the second period. If player 2 deviates in the first period, he can increase his first-period payoff to 4 (by picking Z). But this choice induces player 1 to select B in the second period, where player 2 then best responds with Y. Thus, although a first-period deviation yields an immediate gain of 1, it costs 3 in the second period $(4 - 1)$. This shows that player 2 prefers to behave as prescribed. For his part, player 1 has the incentive to go along with the prescription for play; deviating in either period reduces player 1's payoff. The payoff vector for this subgame perfect equilibrium is enclosed by a pentagon in Figure 22.2(b).

Although the equilibrium construction is a bit complicated, it really is intuitive. Player 2's concern about his reputation and what it implies for his second-period payoff gives him the incentive to forego a short-term gain. If he misbehaves in the first period, his reputation is destroyed and he then suffers in the second period.

Any two-period repeated game can be analyzed as has been done here.[2] Only stage Nash profiles can be played in the second period. However, sometimes reputational equilibria exist whereby the players select non-stage-Nash profiles in the first period. These selections are supported by making the second-period actions contingent on the outcome in the first period (in particular, whether the players cheat or not). The exercises at the end of this chapter will help you better explore the reputation phenomenon.

---

[2]A general analysis of finitely repeated games is reported in J. P. Benoit and V. Krishna, "Finitely Repeated Games," *Econometrica*, 53 (1985): 905–922.

AN INFINITELY REPEATED GAME

Infinitely repeated games are defined by $T = \infty$; that is, the stage game is played each period for an *infinite* number of periods. Although such a game may not seem realistic at first (people do not live forever), infinitely repeated games are useful for modeling some real-world situations. Furthermore, despite the complexity of these games, analysis of their subgame perfect equilibria can actually be quite simple. Consider an infinitely repeated game with *discounting*, whereby the payoffs in the stage game are discounted over time.[3] Let us use δ (a number between 0 and 1) to denote the discount factor for both players. When comparing a payoff received today with a payoff received tomorrow (the next period), we discount tomorrow's payoff by multiplying it by the discount factor. In this way, we say that the stream of payoffs—from today and tomorrow—are "discounted to today." Payoffs obtained two periods from now are discounted by δ^2, payoffs obtained three periods from now are discounted by δ^3, and so on.

For repeated games, we will have to calculate the sum of a stream of discounted payoffs. For example, a player may obtain 1 unit each period for an infinite number of periods. In this case, the sum of his discounted payoff stream is

$$v \equiv 1 + 1\delta + 1\delta^2 + 1\delta^3 + \cdots = 1 + \delta + \delta^2 + \delta^3 + \cdots.$$

We can simplify this expression by noting that

$$\delta + \delta^2 + \delta^3 + \cdots = \delta[1 + \delta + \delta^2 + \delta^3 + \cdots] = \delta v.$$

Therefore, we have

$$v \equiv 1 + \delta v,$$

which means that $v = 1/(1 - \delta)$. In summary,

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

This expression will come in handy. Note that by multiplying both sides by any constant number a, we have

$$a + a\delta + a\delta^2 + a\delta^3 + \cdots = \frac{a}{1-\delta}.$$

The strategies in infinitely repeated games can be exceedingly complex. Recall that in general, a player's strategy is a full description of what action to take at every information set of the player. In a repeated game, there is a different

---

[3]Recall the discussion of discounting in Chapter 19, where the discount factor is used in multistage bargaining games.

FIGURE 22.4
A prisoners' dilemma.

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

information set for every period t and every different history of play from the beginning of the game through period $t - 1$. Thus, a strategy prescribes an action for a player to take conditional on everything that took place in the past.

Fortunately, it is often sufficient to consider just a few types of simple strategies in repeated games. The simplest are those that prescribe stage Nash profiles in each period; as noted in the preceding section, we know these constitute subgame perfect equilibria.

To capture the idea of reputation, we can examine another type of simple strategy called a **trigger strategy**. Trigger strategies specifically refer to two action profiles for the stage game: one profile is called the "cooperative profile," and the other is called the "punishment profile." The punishment profile is assumed to be a stage Nash profile. In a trigger-strategy equilibrium, the players are supposed to play the cooperative profile in each period. However, if one or both of them deviate from the cooperative profile, then they play the punishment profile forever after. In other words, deviating from the cooperative profile destroys a player's reputation and triggers the punishment profile for the rest of the game.

To see how this works, consider the infinitely repeated prisoners' dilemma. The stage game is given in Figure 22.4. There is only one stage Nash equilibrium, (D, D), so we use it as the punishment profile. Let (C, C) be the cooperative profile. Our goal is to understand whether the players have the incentive to play (C, C) each period under the threat that they will revert to (D, D) forever if one or both of them cheat. To be precise, the trigger strategy specifies that the players select (C, C) each period as long as this profile was always played in the past; otherwise, they are to play (D, D). This is sometimes called the *grim-trigger* strategy.

Let us evaluate whether the grim-trigger strategy profile is a subgame perfect equilibrium. Consider the incentives of player i ($i = 1, 2$) from the perspective of period 1. Suppose the other player—called player j—behaves according to the grim trigger. Player i basically has two options. First, she can herself follow the prescription of the grim trigger, which means cooperating as player j does. In this case, player i obtains a payoff of 2 each period, for a discounted total of

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

Second, player i could defect in the first period, which yields an immediate payoff of 3 because player j cooperates in the first period. But player i's defection induces player j to defect in each period thereafter, so then the best that i can do is to keep defecting and get 1 each period. Thus, by defecting in period 1, player i obtains the payoff

$$3 + \delta + \delta^2 + \delta^3 + \cdots = 3 + \delta[1 + \delta + \delta^2 + \delta^3 + \cdots] = 3 + \frac{\delta}{1-\delta}.$$

If

$$\frac{2}{1-\delta} \geq 3 + \frac{\delta}{1-\delta},$$

then player i earns a higher payoff by perpetually cooperating against the grim trigger than by defecting in the first period. Simplifying this inequality yields $\delta \geq 1/2$.

So far, we see that the players have no incentive to cheat in the first period as long as $\delta \geq 1/2$. In fact, the same analysis establishes that the players have no incentive to deviate from the grim trigger in *any* period. For example, suppose the players have cooperated through period $t - 1$. Then, because the game is infinitely repeated, the "continuation game" from period t looks just like the game from period 1, so the analysis starting from period t is exactly the same as the analysis at period 1. Discounting the payoffs to period t, we see that cooperating from period t yields each player $2/(1 - \delta)$. Defecting against the grim trigger leads to the payoff $3 + \delta/(1 - \delta)$. Thus neither player has an incentive to defect in period t if the discount factor exceeds $1/2$.

Analysis of subgame perfection requires us also to look at histories in which one or both players selected D at some point earlier. After such a history, the grim-trigger strategy prescribes play of D. It is clear that as (D, D) is a stage Nash profile, a player has no incentive to deviate from the grim trigger in such a contingency.

In summary, the simple calculation performed in reference to period 1 is enough to establish whether cooperation can be supported in a subgame perfect equilibrium via a reputation mechanism. For the stage game in Figure 22.4, the grim-trigger strategy profile is a subgame perfect Nash equilibrium if and only if $\delta \geq 1/2$.

It is useful to restate the foregoing analysis in terms of *current-period payoffs* and *continuation values*. We can express player i's payoff from any period t as $u_i^t + \delta v_i^{t+1}$, where u_i^t is player i's payoff in period t and v_i^{t+1} is player i's continuation value from the start of period $t + 1$. Note that the latter is multiplied by δ to make it comparable to amounts in period t.

To evaluate whether the grim-trigger strategy profile is a subgame perfect equilibrium, we can appeal to the **one-deviation property** (discussed previously in Chapter 15), which holds for discounted infinitely repeated games. Therefore, we simply check single deviations in two kinds of contingencies: (i) after histories in which (C, C) was always played, and (ii) after histories in which one or both players selected D at some point.

In the latter type of contingency, the grim-trigger strategy profile selects (D, D), which yields a current-period payoff of $u_i^t = 1$ for each player. Further, the continuation value from period $t + 1$ is $v_i^{t+1} = 1/(1 - \delta)$. If a player were to deviate from the grim trigger, his current-period payoff would decrease and his continuation payoff would remain the same, so there is no sense in deviating.

In the former type of contingency, after a history in which (C, C) was always played, the grim-trigger strategy profile yields a current-period payoff of $u_i^t = 2$ and a continuation value of $v_i^{t+1} = 2/(1 - \delta)$ from the start of the next period. If player i were to deviate by selecting D in the current period, his current-period payoff would increase to 3 but his continuation value would decrease to $1/(1 - \delta)$. In other words, the current-period deviation gain is $2 - 1 = 1$, whereas the loss in continuation value is

$$\frac{2}{1-\delta} - \frac{1}{1-\delta} = \frac{1}{1-\delta}.$$

Multiplying the loss by δ to put it in current-period terms, we get the condition under which the players do not want to deviate from grim trigger,

$$1 \leq \delta \frac{1}{1-\delta},$$

which simplifies to $\delta \geq 1/2$.

The grim-trigger analysis applies to any repeated prisoners' dilemma; in fact, it applies to any game in which the players would punish a deviator by reverting to perpetual play of a stage Nash profile. Moreover, the analysis merely requires identifying three stage-game payoffs for each player: the "cooperative payoff" u_i^c, the "deviator's payoff" u_i^d, and the "punishment payoff" u_i^p. These are defined as follows. The value u_i^c is player i's stage-game payoff in the event that the players select the action profile that they hope to support, such as (C, C) in the prisoners' dilemma. The value u_i^d is player i's stage-game payoff in the event that he plays a stage-game best-response to the other player's cooperative action, such as selecting D when the other player selects C in the prisoners' dilemma. Finally, the value u_i^p is player i's stage-game payoff from play of the stage Nash profile, which in the prisoners' dilemma is (D, D).

FIGURE 22.5
Another prisoners' dilemma.

1 \ 2	C	D
C	4, 4	−2, 6
D	6, −2	0, 0

Note that a player who is first to deviate from the grim-trigger profile will get a deviation gain of $u_i^{\mathrm{d}} - u_i^{\mathrm{c}}$ in the current period, but will look forward to a loss of $(u_i^{\mathrm{c}} - u_i^{\mathrm{p}})/(1 - \delta)$ in continuation value starting from the next period. The player prefers not to deviate if

$$u_i^{\mathrm{d}} - u_i^{\mathrm{c}} \leq \frac{\delta}{1-\delta}\,[u_i^{\mathrm{c}} - u_i^{\mathrm{p}}].$$

Solving for δ, this becomes

$$\delta \geq \frac{u_i^{\mathrm{d}} - u_i^{\mathrm{c}}}{u_i^{\mathrm{d}} - u_i^{\mathrm{p}}},$$

which is the necessary and sufficient condition for cooperation to be sustained by the grim-trigger strategy profile as a subgame perfect Nash equilibrium.

For another example, consider the prisoners' dilemma shown in Figure 22.5. For this game we have $u_i^{\mathrm{c}} = 4$, $u_i^{\mathrm{d}} = 6$, and $u_i^{\mathrm{p}} = 0$, so cooperation can be sustained if and only if

$$\delta \geq \frac{6-4}{6-0} = \frac{1}{3}.$$

You can run through calculations like this to double check the examples in the next chapter.

The infinitely repeated game demonstrates that patience—valuing the future—is essential to an effective reputation. When contemplating whether to defect in one period, the players consider the future loss that would result from tarnishing their reputations. Patient players—those with high discount factors—care a lot about payoffs in future periods and therefore they do not want to ruin their reputations for some short-term gain. Thus, there is a sense in which maintaining a reputation is more about the future than the past.[4]

[4] One of the early general analyses of discounted, repeated games may be found in D. Abreu, "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56 (1988): 383–396.

THE EQUILIBRIUM PAYOFF SET WITH LOW DISCOUNTING

The grim-trigger equilibrium discussed in the preceding section is just one of potentially many subgame perfect equilibria in the repeated prisoners' dilemma. We know from the result on page 295 that the following strategy profile also is an equilibrium regardless of the discount factor: play the stage Nash action profile (D, D) in every period, regardless of the history of play. This equilibrium is not very "cooperative," and it yields a low payoff relative to the grim-trigger profile. In this section, I demonstrate that depending on the discount factor, there are many other equilibria exhibiting intermediate amounts of cooperation. The analysis will get a bit technical, but the conclusion at the end is significant.

To develop a picture of the entire set of equilibria in the repeated prisoners' dilemma, consider again the stage game in Figure 22.5. For this stage game, Figure 22.6 depicts the set of feasible stage-game payoffs. This figure also graphs the possible repeated-game payoffs in terms of "average per period," by multiplying the discounted sum payoff by $(1-\delta)$. For example, the point (4, 4) is noted in the picture with a solid circle, which refers to the players obtaining (4, 4) each period in the game [by playing (C, C) each period].[5] The point (6, −2) arises if (D, C) is played each period.

The diamond formed by connecting points (4, 4), (−2, 6), (0, 0), and (6, −2) is important; any payoff vector inside or on the edges of the diamond can be obtained as an average payoff if the players choose the right sequence of

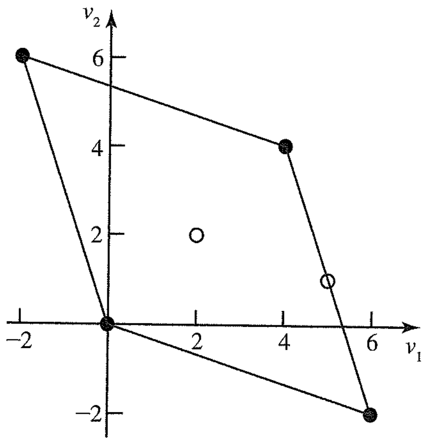


FIGURE 22.6
Possible repeated game payoffs, per period.

[5]Technically speaking, each player's payoff is $4/(1-\delta)$ in this case. In Figure 22.6, I have multiplied this payoff by $(1-\delta)$ to put it in terms of per period.

actions over time. For instance, consider the point (5, 1) designated by an open circle in Figure 22.6. Suppose the players alternate between (C, C) and (D, C) over time, starting with (C, C) in the first period. For player 1, this sequence of actions yields a discounted payoff of

$$4 + 6\delta + 4\delta^2 + 6\delta^3 + \cdots.$$

Factoring terms, this expression simplifies to

$$4[1 + \delta^2 + \delta^4 + \cdots] + 6\delta[1 + \delta^2 + \delta^4 + \cdots] = \frac{4}{1 - \delta^2} + \frac{6\delta}{1 - \delta^2}.$$

Here I have used the same method as that put to use earlier to calculate the sums.[6] Recognizing that $1 - \delta^2 = (1 + \delta)(1 - \delta)$, we can write player 1's payoff as

$$\frac{4 + 6\delta}{(1 - \delta)(1 + \delta)}.$$

Multiplying by $(1 - \delta)$ puts this in terms of average per period:

$$\frac{4 + 6\delta}{1 + \delta}.$$

Likewise, player 2's per-period average is

$$\frac{4 - 2\delta}{1 + \delta}.$$

Note that if δ is close to 1, then this average payoff vector is arbitrarily close to (5, 1). Just plug in $\delta = 1$ to see this. You can try other examples to convince yourself that the diamond in Figure 22.6 represents the set of average per-period payoffs that can arise in the repeated game. For instance, determine how an average payoff of (2, 2) can be obtained.

With the set of feasible repeated-game payoffs in mind, we can determine whether or not any particular payoff vector can be supported as the result of a subgame perfect equilibrium. For an example, focus on the per-period average vector (5, 1) and consider the following "modified grim-trigger strategy": The players are instructed to alternate between (C, C) and (D, C) over time, starting with (C, C) in the first period. If either or both players has deviated from this prescription in the past, the players are supposed to revert to the stage Nash profile (D, D) forever. To determine whether this strategy profile is an equilibrium, we must compare each player's short-term gain from deviating to the punishment.

[6]Letting $v = 1 + \delta^2 + \delta^4 + \cdots$, we have $v = 1 + \delta^2 v$. Solving, we get $v = 1/(1 - \delta^2)$.

Let us begin with the incentives of player 2. First note that if the players conform to the modified grim trigger, then, starting from any odd-numbered period [in which players select (C, C)], player 2's payoff is

$$4 - 2\delta + 4\delta^2 - 2\delta^3 + \cdots = \frac{4 - 2\delta}{1 - \delta^2}.$$

Starting from any even-numbered period, player 2's payoff is

$$-2 + 4\delta - 2\delta^2 + 4\delta^3 + \cdots = \frac{-2 + 4\delta}{1 - \delta^2}.$$

We need to check whether player 2 has an incentive to cheat in either an odd-numbered or an even-numbered period. Note that in each odd-numbered period, the players are supposed to select (C, C). If player 2 cheats (by defecting), then he obtains a short-term gain of $6 - 4 = 2$; however, his continuation value starting in the next period—an even-numbered period—will be 0 (because the players defect thereafter) rather than $(-2 + 4\delta)/(1 - \delta^2)$. If we discount this future loss to the period in which player 2 cheats, the payoff is $\delta(-2 + 4\delta)/(1 - \delta^2)$. Thus, player 2 prefers to cooperate in odd-numbered periods if and only if the long-term loss of cheating outweighs the short-term gain:

$$\frac{\delta(-2 + 4\delta)}{(1 - \delta^2)} \geq 2. \tag{1}$$

Repeating the calculation for even-numbered periods yields the following inequality:

$$\frac{\delta(4 - 2\delta)}{(1 - \delta^2)} \geq 2. \tag{2}$$

The short-term gain of 2 here is due to player 2 obtaining 0 rather than -2.

In summary, assuming player 1 plays according to the modified grim trigger, player 2 wishes to conform if and only if both inequalities 1 and 2 are satisfied. You can verify with a bit of algebraic manipulation that the first simplifies to $\delta \geq (1 + \sqrt{13})/6$ and the second simplifies to $\delta \geq 1/2$. Because $(1 + \sqrt{13})/6 > 1/2$, the first inequality is more stringent than is the second. Thus, player 2 cooperates as long as $\delta \geq (1 + \sqrt{13})/6$. You can perform the same kind of analysis to find that player 1 will conform to the modified grim trigger as long as $\delta \geq (-3 + \sqrt{21})/6$. Because

$$\frac{1 + \sqrt{13}}{6} > \frac{-3 + \sqrt{21}}{6},$$

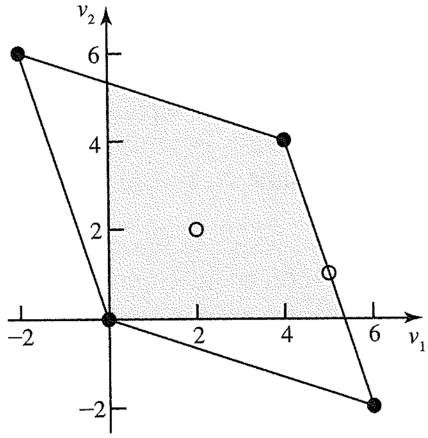


FIGURE 22.7
Equilibrium per-period payoffs for large δ.

we conclude that the modified grim-trigger profile is a subgame perfect equilibrium if and only if $\delta \geq (1 + \sqrt{13})/6$. Furthermore, as already noted, the payoff of this equilibrium is close to (5, 1) if δ is close to 1.

I realize that you might find this analysis complicated. But putting the details of the arguments aside, recognize the general conclusion. I showed that a point on one of the edges of the diamond in Figure 22.6 can be supported as an equilibrium average per-period payoff as long as the players are patient enough. In fact, *any* point on the edge or interior of the diamond can be so supported, as long as two conditions hold: (1) each player obtains more than 0, and (2) the discount factor is close enough to 1. Figure 22.7 depicts the set of equilibrium payoffs. Here is the result stated for general repeated games:

> **Result:** Consider any infinitely repeated game. Suppose there is a stage Nash profile that yields payoff vector w (w_i for player i, $i = 1, 2, \ldots, n$). Let v be any feasible average per-period payoff such that $v_i > w_i$ for each player i. The vector v can be supported arbitrarily closely by a subgame perfect Nash equilibrium if δ is close enough to 1.

In other words, with the use of trigger strategies, almost any repeated-game payoff can be achieved in equilibrium with patient players.[7]

[7]Game theorists call this the *folk theorem* because it was thought to have been a part of the profession's conventional wisdom before versions of the result were formally proved. A general treatment appears in D. Fudenberg and E. Maskin, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54 (1986): 533–554. For his pioneering analysis of repeated games and the folk theorem (among other theoretical contributions), Robert Aumann was awarded the Nobel Prize in 2005.

GUIDED EXERCISE

Problem: Persons 1 and 2 are forming a firm. The value of their relationship depends on the effort that each expends. Suppose that person i's utility from the relationship is $x_j^2 + x_j - x_i^2$, where x_i is person i's effort and x_j is the effort of the other person ($i = 1, 2$). Assume $x_1, x_2 \in [0, 100]$.

(a) Compute the partners' best-response functions and find the Nash equilibrium of this game. Is the Nash equilibrium efficient?
(b) Now suppose that the partners interact over time, which we model with the infinitely repeated version of the game. Let δ denote the discount factor of the players. Under what conditions can the partners sustain some positive effort level $k = x_1 = x_2$ over time?
(c) Comment on how the maximum sustainable effort depends on the partners' patience.

Solution:

(a) To find the best-response functions, fix player j's effort x_j and consider how player i's effort level x_i affects his payoff. Observe that player i's payoff is strictly decreasing in x_i and so player i's best response is to select the lowest effort $x_i = 0$. The only Nash equilibrium of this game is the profile (0, 0); that is, each player selects the lowest effort level. The equilibrium is not efficient. To see this, note that if both players select effort level $k > 0$, then they each get the payoff $k^2 + k - k^2 = k$, which exceeds the payoff of 0 that they get with profile (0, 0). In other words, profile (0, 0) is less efficient than profile (k, k).
(b) Each player can guarantee himself a payoff of 0 by selecting $x_i = 0$ in every period. Thus, repeated play of the stage Nash profile, which yields a payoff of 0, is the worst punishment that can be levied against a player. With this in mind, consider the grim-trigger strategy profile in which each player selects effort level k in each period, as long as both players had done so in the past. Otherwise, the players revert to the stage Nash profile. When this strategy profile is played, then each player gets k in each period. The discounted sum over an infinite number of periods is $k/(1 - \delta)$. Note that the best way for a player to deviate is to select 0 effort rather than k, which yields a payoff of $k^2 + k$ in the period of the deviation (as the other player selects k). Following the deviation, the players will revert to the stage Nash profile, which gives the deviator a payoff of 0 in all future periods. Thus, for the grim trigger to be an equilibrium, we need $k/(1 - \delta) \geq k^2 + k$, which simplifies to $\delta/(1 - \delta) \geq k$. Rearranging yields $\delta \geq k/(1 + k)$.

(c) The maximal sustainable effort is $\delta/(1 - \delta)$, which is increasing in δ. In other words, more patient players can sustain higher levels of effort.

EXERCISES

1. Consider a repeated game in which the stage game in the following figure is played in each of two periods and there is no discounting.

1 \ 2	L	M	R
U	8, 8	0, 9	0, 0
C	9, 0	0, 0	3, 1
D	0, 0	1, 3	3, 3

Fully describe a subgame perfect equilibrium in which the players select (U, L) in the first period.

2. Find conditions on the discount factor under which cooperation can be supported in the infinitely repeated games with the following stage games.

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

(a)

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

(b)

1 \ 2	C	D
C	3, 2	0, 1
D	7, 0	2, 1

(c)

Use the grim-trigger strategy profile.

3. Consider the following stage game.

1 \ 2	X	Y
A	5, 6	0, 0
B	8, 2	2, 2

(a) Find and report all of the (pure-strategy) Nash equilibria of this game.

(b) Consider the two-period repeated game in which this stage game is played twice and the repeated-game payoffs are simply the sum of the payoffs in each of the two periods. Is there a subgame perfect equilibrium of this repeated game in which (A, X) is played in the first period? If so, fully describe the equilibrium. If not, explain why.

4. If its stage game has exactly one Nash equilibrium, how many subgame perfect equilibria does a two-period, repeated game have? Explain. Would your answer change if there were T periods, where T is any finite integer?

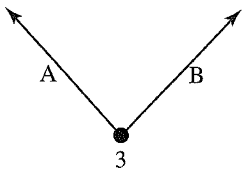
5. Consider the infinitely repeated game where the stage game is the matrix in Exercise 2(c). Under what conditions is there a subgame perfect equilibrium in which the players alternate between (C, C) and (C, D), starting with (C, C) in the first period? Under what conditions is there a subgame perfect equilibrium in which the players alternate between (C, C) and (D, D), starting with (C, C) in the first period? (Use modified trigger strategies.)

6. Which is more important to achieving cooperation through a reputation, a long history together or a long horizon ahead?

7. Consider the following three-player game.

1 \ 2	L	R
U	2, 2, 0	5, 5, 5
D	8, 6, 8	0, 7, 4

1 \ 2	L	R
U	4, 4, 1	4, 2, 8
D	0, 2, 9	4, 2, 5



The players make their choices simultaneously and independently. The payoffs are listed in order of the player numbers.

(a) Find the (pure-strategy) Nash equilibria of this game.

(b) Consider the two-period repeated game in which this stage game is played twice and the repeated game payoffs are simply the sum of the payoffs in the two periods. Compute and report all of the subgame perfect equilibria of this repeated game. List the set of subgame perfect equilibrium payoffs.

8. Consider an infinite-period repeated game in which a "long-run player" faces a sequence of "short-run" opponents. Formally, player 1 plays the stage game with a different player 2 in each successive period. Denote by 2^t the player who plays the role of player 2 in the stage game in period t. Assume that all players observe the history of play. Let δ denote the discount factor of player 1. Note that such a game has an infinite number of players.

(a) In any subgame perfect equilibrium, what must be true about the behavior of player 2^t with respect to the action selected by player 1 in period t?

(b) Give an example of a stage-game and subgame perfect equilibrium where the players select an action profile in the stage game that is not a stage Nash equilibrium.

(c) Show by example that a greater range of behavior can be supported when both players are long-run players than when only player 1 is a long-run player.

9. Consider the following "war of attrition" game. Interaction between players 1 and 2 takes place over discrete periods of time, starting in period 1. In each period, players choose between "stop" (S) and "continue" (C) and they receive payoffs given by the following stage-game matrix:

1 \ 2	S	C
S	x, x	0, 10
C	10, 0	−1, −1

The length of the game depends on the players' behavior. Specifically, if one or both players select S in a period, then the game ends at the end of this period. Otherwise, the game continues into the next period. Suppose the players discount payoffs between periods according to discount factor δ. Assume $x < 10$.

(a) Show that this game has a subgame perfect equilibrium in which player 1 chooses S and player 2 chooses C in the first period. Note that in such an equilibrium, the game ends at the end of period 1.

(b) Assume $x = 0$. Compute the symmetric equilibrium of this game. (Hint: In each period, the players randomize between C and S. Let α denote the probability that each player selects S in a given period.)

(c) Write an expression for the symmetric equilibrium value of α for the case in which x is not equal to 0.

10. Consider a repeated game between a supplier (player 1) and a buyer (player 2). These two parties interact over an infinite number of periods. In each period, player 1 chooses a quality level $q \in [0, 5]$ at cost q. Simultaneously, player 2 decides whether to purchase the good at a fixed price of 6. If player 2 purchases, then the stage-game payoffs are $6 - q$ for player 1 and $2q - 6$ for player 2. Here, player 2 is getting a benefit of $2q$. If player 2 does not purchase, then the stage-game payoffs are $-q$ for player 1 and 0 for player 2. Suppose that both players have discount factor δ.

(a) Calculate the efficient quality level under the assumption that transfers are possible (so you should look at the sum of payoffs).

(b) For sufficiently large δ, does this game have a subgame perfect Nash equilibrium that yields the efficient outcome in each period? If so, describe the equilibrium strategies and determine how large δ must be for this equilibrium to exist.

11. This is an extension of the previous exercise. Consider the following stage game between a manager (also called the "Principal") and a worker (the "Agent"). Let the manager be player 1 and the worker be player 2. Simultaneously, the manager chooses a bonus payment $p \in [0, \infty)$ and the worker chooses an effort level $a \in [0, \infty)$. The stage-game payoffs are $u_1(p, a) = 4a - p$ and $u_2(p, a) = p - a^2$.

(a) Determine the efficient effort level for the worker.

(b) Find the Nash equilibrium of the stage game.

(c) Suppose the stage game is to be played twice (a two-period repeated game) and there is no discounting. Find all of the subgame perfect equilibria.

(d) Suppose the stage game is to be played infinitely many times in succession (an infinitely repeated game) and assume that the players share the discount factor $\delta < 1$. Find conditions on the discount factor under which there is a subgame perfect equilibrium featuring selection of the efficient effort level in each period (on the equilibrium path).

12. Consider the infinitely repeated prisoners' dilemma and recall the definition of the grim-trigger strategy. Here is the definition of another simple strategy called *Tit-for-tat:* Select C in the first period; in each period thereafter, choose the action that the opponent selected in the previous period.[8] Is the tit-for-tat strategy profile a Nash equilibrium of the repeated game for discount factors close to one? Is this strategy profile a subgame perfect Nash equilibrium of the repeated game for discount factors close to one? Explain.

[8]For more on tit-for-tat, you might look at R. Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).

23 COLLUSION, TRADE AGREEMENTS, AND GOODWILL

In this chapter, I sketch three applications of repeated-game theory. Two of them elaborate on analysis presented in Part II of this book. In particular, to study collusion between firms, I use a repeated version of the Cournot duopoly model; discussion of the enforcement of international trade agreements utilizes a similar repeated game.

DYNAMIC OLIGOPOLY AND COLLUSION

Consider the Cournot duopoly model in Chapter 10, with two firms that each produce at zero cost (which I assume just to make the computations easy), and suppose the market price is given by $p = 1 - q_1 - q_2$. Firm i, which produces q_i, obtains a payoff of $(1 - q_i - q_j)q_i$. Note that the Nash equilibrium of this game is $q_1 = q_2 = 1/3$, yielding a payoff of $1/9$ for each firm. As noted in Chapter 10, this outcome is inefficient from the firms' point of view; they would both be better off if they shared the monopoly level of output by each producing $1/4$. Sharing the monopoly output yields each firm a payoff of $1/8$, which is greater than the Nash equilibrium payoff of $1/9$.[1] In the static game, therefore, the firms would like to collude to set $q_1 = q_2 = 1/4$, but this strategy profile cannot be sustained because it is not an equilibrium.

In most industries, firms do not interact in just a single point in time. They interact every day, potentially forever. To model firms' ongoing interaction, we can examine an infinitely repeated version of the Cournot duopoly, where the stage game is defined as the Cournot game described in the preceding paragraph. Analysis of the infinitely repeated game demonstrates that collusion can be sustained in equilibrium, using the reputation mechanism. In particular, let us evaluate the following grim-trigger strategy profile: Each firm is prescribed to select $1/4$ in each period, as long as both firms did so in the past; if one or both players deviates, the firms are supposed to play the stage Nash profile $(1/3, 1/3)$ forever after.

[1] If numbers $1/8$ and $1/9$ seem insignificant, think of q_i as millions of units, the price in dollars, and the payoff therefore in millions of dollars.

10. Consider a repeated game between a supplier (player 1) and a buyer (player 2). These two parties interact over an infinite number of periods. In each period, player 1 chooses a quality level $q \in [0, 5]$ at cost q. Simultaneously, player 2 decides whether to purchase the good at a fixed price of 6. If player 2 purchases, then the stage-game payoffs are $6 - q$ for player 1 and $2q - 6$ for player 2. Here, player 2 is getting a benefit of $2q$. If player 2 does not purchase, then the stage-game payoffs are $-q$ for player 1 and 0 for player 2. Suppose that both players have discount factor δ.

(a) Calculate the efficient quality level under the assumption that transfers are possible (so you should look at the sum of payoffs).

(b) For sufficiently large δ, does this game have a subgame perfect Nash equilibrium that yields the efficient outcome in each period? If so, describe the equilibrium strategies and determine how large δ must be for this equilibrium to exist.

11. This is an extension of the previous exercise. Consider the following stage game between a manager (also called the "Principal") and a worker (the "Agent"). Let the manager be player 1 and the worker be player 2. Simultaneously, the manager chooses a bonus payment $p \in [0, \infty)$ and the worker chooses an effort level $a \in [0, \infty)$. The stage-game payoffs are $u_1(p, a) = 4a - p$ and $u_2(p, a) = p - a^2$.

(a) Determine the efficient effort level for the worker.

(b) Find the Nash equilibrium of the stage game.

(c) Suppose the stage game is to be played twice (a two-period repeated game) and there is no discounting. Find all of the subgame perfect equilibria.

(d) Suppose the stage game is to be played infinitely many times in succession (an infinitely repeated game) and assume that the players share the discount factor $\delta < 1$. Find conditions on the discount factor under which there is a subgame perfect equilibrium featuring selection of the efficient effort level in each period (on the equilibrium path).

12. Consider the infinitely repeated prisoners' dilemma and recall the definition of the grim-trigger strategy. Here is the definition of another simple strategy called *Tit-for-tat:* Select C in the first period; in each period thereafter, choose the action that the opponent selected in the previous period.[8] Is the tit-for-tat strategy profile a Nash equilibrium of the repeated game for discount factors close to one? Is this strategy profile a subgame perfect Nash equilibrium of the repeated game for discount factors close to one? Explain.

[8]For more on tit-for-tat, you might look at R. Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).