GAMES OF STRATEGY

THIRD EDITION

Avinash Dixit
Princeton University

Susan Skeath
Wellesley College

David Reiley
University of Arizona and Yahoo! Research

W. W. Norton & Company
New York • London

(e) Explain why the children would want to make a threat in the first place, and suggest a way in which they might make their threatened action credible.

U5. Answer the questions in Exercise S5 for the following situations:

(a) The students at your university or college want to prevent the administration from raising tuition.

(b) Most participants, as well as outsiders, want to achieve a durable peace in Afghanistan, Iraq, Israel, and Palestine.

(c) Nearly all nations of the world want Iran to shut down its nuclear program.

U6. Write a brief description of a game in which you have participated, entailing strategic moves such as a commitment, threat, or promise and paying special attention to the essential aspect of credibility. Provide an illustration of the game if possible, and explain why the game that you describe ended as it did. Did the players use sound strategic thinking in making their choices?

11

The Prisoners' Dilemma and Repeated Games

In this chapter, we continue our study of broad classes of games with an analysis of the prisoners' dilemma game. It is probably *the* classic example of the theory of strategy and its implications for predicting the behavior of game players, and most people who learn only a little bit of game theory learn about it. Even people who know *no* game theory may know the basic story behind this game or they may have at least heard that it exists. The prisoners' dilemma is a game in which each player has a dominant strategy, but the equilibrium that arises when all players use their dominant strategies provides a worse outcome for every player than would arise if they all used their dominated strategies instead. The paradoxical nature of this equilibrium outcome leads to several more complex questions about the nature of the interactions that only a more thorough analysis can hope to answer. The purpose of this chapter is to provide that additional thoroughness.

We already considered the prisoners' dilemma in Section 3 of Chapter 4. There we took note of the curious nature of the equilibrium that is actually a "bad" outcome for the players. The "prisoners" can find another outcome that both prefer to the equilibrium outcome, but they find it difficult to bring about. The focus of this chapter is the potential for achieving that better outcome. That is, we consider whether and how the players in a prisoners' dilemma can attain and sustain their mutually beneficial cooperative outcome, overcoming their separate incentives to defect for individual gain. We first review the standard prisoners' dilemma game and then develop four categories of solutions. The first and most important method of solution consists of repetition of the standard one-shot game. The general theory of repeated games was the contribution for which Robert Aumann was awarded the 2005 Nobel Prize in Economics (jointly with Thomas Schelling). As usual at this introductory level, we look at a few simple examples of this general theory. Two other potential solutions rely on penalty (or reward) schemes and on the role of leadership. The fourth incorporates asymmetric information into a finitely repeated dilemma game. As we consider each potential solution, the importance of the costs of defecting and the benefits of cooperation will become clear.

This chapter concludes with a discussion of some of the experimental evidence regarding the prisoners' dilemma as well as several examples of actual dilemmas in action. Experiments generally put live players in a variety of prisoners' dilemma–type games and show some perplexing as well as some more predictable behavior; experiments conducted with the use of computer simulations yield additional interesting outcomes. Our examples of real-world dilemmas that end the chapter are provided to give a sense of the diversity of situations in which prisoners' dilemmas arise and to show how, in at least one case, players may be able to create their own solution to the dilemma.

1 THE BASIC GAME (REVIEW)

Before we consider methods for avoiding the "bad" outcome in the prisoners' dilemma, we briefly review the basics of the game. Recall our example from Chapter 4 of the husband and wife suspected of murder. Each is interrogated separately and can choose to confess to the crime or to deny any involvement. The payoff matrix that they face was originally presented as Figure 4.4 and is reproduced here as Figure 11.1. The numbers shown indicate years in jail; therefore low numbers are better for both players.

Both players here have a dominant strategy. Each does better to confess, regardless of what the other player does. The equilibrium outcome entails both players deciding to confess and each getting 10 years in jail. If they both had

		WIFE	
		Confess (Defect)	Deny (Cooperate)
HUSBAND	Confess (Defect)	10 yr, 10 yr	1 yr, 25 yr
	Deny (Cooperate)	25 yr, 1 yr	3 yr, 3 yr

FIGURE 11.1 Payoffs for the Standard Prisoners' Dilemma

chosen to deny any involvement, however, they would have been better off, with only 3 years of jail time to serve.

In any prisoners' dilemma game, there is always a *cooperative strategy* and a *cheating or defecting strategy*. In Figure 11.1, Deny is the cooperative strategy; both players using that strategy yields the best outcome for the players. Confess is the cheating or defecting strategy; when the players do not cooperate with one another, they choose to Confess in the hope of attaining individual gain at the rival's expense. Thus, players in a prisoners' dilemma can always be labeled, according to their choice of strategy, as either *defectors* or *cooperators*. We will use this labeling system throughout the discussion of potential solutions to the dilemma.

We want to emphasize that, although we speak of a cooperative strategy, the prisoners' dilemma game is noncooperative in the sense explained in Chapter 2—namely, the players make their decisions and implement their choices individually. If the two players could discuss, choose, and play their strategies jointly—as if, for example, the prisoners were in the same room and could give a joint answer to the question of whether they were both going to confess—there would be no difficulty about their achieving the outcome that both prefer. The essence of the questions of whether, when, and how a prisoners' dilemma can be resolved is the difficulty of achieving a cooperative (jointly preferred) outcome through noncooperative (individual) actions.

2 SOLUTIONS I: REPETITION

Of all the mechanisms that can sustain cooperation in the prisoners' dilemma, the best known and the most natural is **repeated play** of the game. Repeated or ongoing relationships between players imply special characteristics for the games that they play against one another. In the prisoners' dilemma, this result plays out in the fact that each player fears that one instance of defecting will lead to a collapse of cooperation in the future. If the value of future cooperation is large and exceeds what can be gained in the short term by defecting, then the long-term individual interests of the players can automatically and tacitly keep them from defecting, without the need for any additional punishments or enforcement by third parties.

We consider the meal-pricing dilemma faced by the two restaurants, Xavier's Tapas and Yvonne's Bistro, introduced in Chapter 5. For our purposes here, we have chosen to simplify that game by supposing that only two choices of price are available: the jointly best (collusive) price of \$26 or the Nash equilibrium price of \$20. The payoffs (profits measured in hundreds of dollars per month) for each restaurant can be calculated by using the quantity (demand) functions in Section 1.A of Chapter 5; these payoffs are shown in Figure 11.2. As in any

		YVONNE'S BISTRO	
		20 (Defect)	26 (Cooperate)
XAVIER'S TAPAS	20 (Defect)	288, 288	360, 216
	26 (Cooperate)	216, 360	324, 324

FIGURE 11.2 Prisoners' Dilemma of Pricing ($100s per month)

prisoners' dilemma, each store has a dominant strategy to defect and price its meals at $20, although both stores would prefer the outcome in which each cooperates and charges the higher price of $26 per meal.

Let us start our analysis by supposing that the two restaurants are initially in the cooperative mode, each charging the higher price of $26. If one restaurant—say, Xavier's—deviates from this pricing strategy, it can increase its profit from 324 to 360 (from $32,400 to $36,000) for one month. But then cooperation has dissolved and Xavier's rival, Yvonne's, will see no reason to cooperate from then on. Once cooperation has broken down, presumably permanently, the profit for Xavier's is 288 each month instead of the 324 it would have been if Xavier's had never defected in the first place. By gaining 36 ($3,600) in one month of defecting, Xavier's gives up 36 ($3,600) each month thereafter by destroying cooperation. Even if the relationship lasts as little as three months, it seems that defecting is not in Xavier's best interest. A similar argument can be made for Yvonne's. Thus, if the two restaurants competed on a regular basis for at least three months, it seems that we might see cooperative behavior and high prices rather than the defecting behavior and low prices predicted by theory for the one-shot game.

A. Finite Repetition

But the solution of the dilemma is not actually that simple. What if the relationship did last exactly three months? Then strategic restaurants would want to analyze the full three-month game and choose their optimal pricing strategies. Each would use rollback to determine what price to charge each month. Starting their analyses with the third month, they would realize that, at that point, there was no future relationship to consider. Each restaurant would find that it had a dominant strategy to defect. Given that, there is effectively no future to consider in the second month either. Each player knows that there will be mutual defecting in the third month, and therefore both will defect in the second month; defecting is the dominant strategy in month 2 also. Then the same argument applies to the first month as well. Knowing that both will defect in months 2 and 3 anyway, there is no future value of cooperation

in the first month. Both players defect right from the start, and the dilemma is alive and well.

This result is very general. As long as the relationship between the two players in a prisoners' dilemma game lasts a fixed and known length of time, the dominant-strategy equilibrium with defecting should prevail in the last period of play. When the players arrive at the end of the game, there is never any value to continued cooperation, and so they defect. Then rollback predicts mutual defecting all the way back to the very first play. However, in practice, players in finitely repeated prisoners' dilemma games show a lot of cooperation; more on this to come.

B. Infinite Repetition

Analysis of the finitely repeated prisoners' dilemma shows that even repetition of the game cannot guarantee the players a solution to their dilemma. But what would happen if the relationship did not have a predetermined length? What if the two restaurants expected to continue competing with one another indefinitely? Then our analysis must change to incorporate this new aspect of their interaction, and we will see that the incentives of the players change also.

In repeated games of any kind, the sequential nature of the relationship means that players can adopt strategies that depend on behavior in preceding plays of the games. Such strategies are known as **contingent strategies,** and several specific examples are used frequently in the theory of repeated games. Most contingent strategies are **trigger strategies.** A player using a trigger strategy plays cooperatively as long as her rival(s) do so, but any defection on their part "triggers" a period of **punishment,** of specified length, in which she plays noncooperatively in response. Two of the best-known trigger strategies are the grim strategy and tit-for-tat. The **grim strategy** entails cooperating with your rival until such time as she defects from cooperation; once a defection has occurred, you punish your rival (by choosing the Defect strategy) on every play for the rest of the game.[1] **Tit-for-tat (TFT)** is not so harshly unforgiving as the grim strategy and is famous (or infamous) for its ability to solve the prisoners' dilemma without requiring permanent punishment. Playing TFT means choosing, in any specified period of play, the action chosen by your rival in the preceding period of play. Thus, when playing TFT, you cooperate with your rival if she cooperated during the most recent play of the game and defect (as punishment) if your rival defected. The punishment phase lasts only as long as your rival continues to defect; you will return to cooperation one period after she chooses to do so.

[1] Defecting as retaliation under the requirements of a trigger strategy is often termed *punishing* to distinguish it from the original decision to deviate from cooperation.

Let us consider how play might proceed in the repeated restaurant pricing game if one of the players uses the contingent strategy tit-for-tat. We have already seen that if Xavier's Tapas defects one month, it could add 36 to its profits (360 instead of 324). But if Xavier's rival is playing TFT, then such defecting would induce Yvonne's Bistro to punish Xavier's the next month in retaliation. At that point, Xavier's has two choices. One option is to continue to defect by pricing at \$20, and to endure Yvonne's continued punishment according to TFT; in this case, Xavier's loses 36 (288 rather than 324) for every month thereafter in the foreseeable future. This option appears quite costly. But Xavier's *could* get back to cooperation, too, if it so desired. By reverting to the cooperative price of \$26 after one month's defection, Xavier's would incur only one month's punishment from Yvonne's. During that month, Xavier's would suffer a loss in profit of 108 (216 rather than the 324 that would have been earned without any defection). In the second month after Xavier's defection, both restaurants could be back at the cooperative price earning 324 each month. This one-time defection yields an extra 36 in profit but costs an additional 108 during the punishment, also apparently quite costly to Xavier's.

It is important to realize here, however, that Xavier's extra \$36 from defecting is gained in the first month. Its losses are ceded in the future. Therefore the relative importance of the two depends on the relative importance of the present versus the future. Here, because payoffs are calculated in dollar terms, an objective comparison can be made. Generally, money (or profit) that is earned today is better than money that is earned later because, even if you do not need (or want) the money until later, you can invest it now and earn a return on it until you need it. So Xavier's should be able to calculate whether it is worthwhile to defect, on the basis of the total rate of return on its investment (including capital gains and/or dividends and/or interest, depending on the type of investment). We use the symbol r to denote this rate of return. Thus one dollar invested generates r dollars of interest and/or dividends and/or capital gains, or 100 dollars generate $100r$, therefore the rate of return is sometimes also said to be $100r\%$.

Note that we can calculate whether it is in Xavier's interest to defect because the firms' payoffs are given in dollar terms, rather than as simple ratings of outcomes, as in some of the games in earlier chapters (the street-garden game in Chapters 3 and 6, for example). This means that payoff values in different cells are directly comparable; a payoff of 4 (dollars) is twice as good as a payoff of 2 (dollars) here, whereas a payoff of 4 is not necessarily exactly twice as good as a payoff of 2 in any two-by-two game in which the four possible outcomes are ranked from 1 (worst) to 4 (best). As long as the payoffs to the players are given in measurable units, we can calculate whether defecting in a prisoners' dilemma game is worthwhile.

I. IS IT WORTHWHILE TO DEFECT ONLY ONCE AGAINST A RIVAL PLAYING TFT? One of Xavier's options when playing repeatedly against a rival using TFT is to defect just once from a cooperative outcome and then to return to cooperating. This particular strategy gains the restaurant 36 in the first month (the month during which it defects) but loses it 108 in the second month. By the third month, cooperation is restored. Is defecting for only one month worth it?

We cannot directly compare the 36 gained in the first month with the 108 lost in the second month, because the additional money value of time must be incorporated into the calculation. That is, we need a way to determine how much the 108 lost in the second month is worth during the first month. Then we can compare that number with 36 to see whether defecting once is worthwhile. What we are looking for is the **present value (PV)** of 108, or how much in profit earned this month (in the present) is equivalent to (has the same value as) the 108 earned next month. We need to determine the number of dollars earned this month that, with interest, would give us 108 next month; we call that number PV, the present value of 108.

Given that the (monthly) total rate of return is r, getting PV this month and investing it until next month yields a total next month of PV + rPV, where the first term is the principal being paid back and the second term is the return (interest or dividend or capital gain). When the total is exactly 108, then PV equals the present value of 108. Setting PV + rPV = 108 yields a solution for PV:

$$PV = \frac{108}{1+r}.$$

For any value of r, we can now determine the exact number of dollars that, earned this month, would be worth 108 next month.

From the perspective of Xavier's Tapas, the question remains whether the gain of 36 this month is offset by the loss of 108 next month. The answer depends on the value of PV. Xavier's must compare the gain of 36 with the PV of the loss of 108. To defect once (and then return to cooperation) is worthwhile only if $36 > 108/(1 + r)$. This is the same as saying that defecting once is beneficial only if $36(1 + r) > 108$, which reduces to $r > 2$. Thus Xavier's should choose to defect once against a rival playing TFT only if the monthly total rate of return exceeds 200%. This outcome is very unlikely; for example, prime lending rates rarely exceed 12% per year. This translates into a monthly interest rate of no more than 1% (compounded annually, not monthly), well below the 200% just calculated. Here, it is better for Xavier's to continue cooperating than to try a single instance of defecting when Yvonne's is playing TFT.

II. IS IT WORTHWHILE TO DEFECT FOREVER AGAINST A RIVAL PLAYING TFT? What about the possibility of defecting once and then continuing to defect forever? This second option of

Xavier's gains the restaurant 36 in the first month but loses it 36 in every month thereafter into the future if the rival restaurant plays TFT. To determine whether such a strategy is in Xavier's best interest again depends on the present value of the losses incurred. But this time the losses are incurred over an **infinite horizon** of future months of competition.

We need to figure out the present value of all of the 36s that are lost in future months, add them all up, and compare them with the 36 gained during the month of defecting. The PV of the 36 lost during the first month of punishment and continued defecting on Xavier's part is just $36/(1 + r)$; the calculation is identical with that used in Section 2.B.I to find that the PV of 108 was $108/(1 + r)$. For the next month, the PV must be the dollar amount needed this month that, with two months of **compound interest,** would yield 36 in two months. If the PV is invested now, then in one month the investor would have that principal amount plus a return of rPV, for a total of PV + rPV, as before; leaving this total amount invested for the second month means that at the end of two months, the investor has the amount invested at the beginning of the second month (PV + rPV) plus the return on that amount, which would be r(PV + rPV). The PV of the 36 lost two months from now must then solve the equation: PV + rPV + r(PV + rPV) = 36. Working out the value of PV here yields PV$(1 + r)^2 = 36$, or PV = $36/(1 + r)^2$. You should see a pattern developing. The PV of the 36 lost in the third month of continued defecting is $36/(1 + r)^3$, and the PV of the 36 lost in the fourth month is $36/(1 + r)^4$. In fact, the PV of the 36 lost in the nth month of continued defecting is just $36/(1 + r)^n$. Xavier's loses an infinite sum of 36s, and the PV of each of them gets smaller each month.

More precisely, Xavier's loses the sum, from $n = 1$ to $n = \infty$ (where n labels the months of continued defecting after the initial month), of $36/(1 + r)^n$. Mathematically, it is written as the sum of an infinite number of terms:[2]

$$36/(1 + r) + 36/(1 + r)^2 + 36/(1 + r)^3 + 36/(1 + r)^4 + \cdots .$$

Because r is a rate of return and presumably a positive number, the ratio of $1/(1 + r)$ will be less than 1; this ratio is generally called the **discount factor** and is referred to by the Greek letter δ. With $\delta = 1/(1 + r) < 1$, the mathematical rule for infinite sums tells us that this sum converges to a specific value, in this case $36/r$.

It is now possible to determine whether Xavier's Tapas will choose to defect forever. The restaurant compares its gain of 36 with the PV of all the lost 36s, or $36/r$. Then it defects forever only if $36 > 36/r$, or $r > 1$; defecting forever is beneficial in this particular game only if the monthly rate of return exceeds 100%, an unlikely event. Thus we would not expect Xavier's to defect against a cooperative rival when both are playing tit-for-tat. When both Yvonne's Bistro and Xavier's

[2]The Appendix to this chapter contains a detailed discussion of the solution of infinite sums.

Tapas play TFT, the cooperative outcome in which both price high is a Nash equilibrium of the game. Both playing TFT is a Nash equilibrium, and use of this contingent strategy solves the prisoners' dilemma for the two restaurants.

Remember that tit-for-tat is only one of many trigger strategies that players could use in repeated prisoners' dilemmas. And it is one of the "nicer" ones. Thus if TFT can be used to solve the dilemma for the two restaurants, other, harsher trigger strategies should be able to do the same. The grim strategy, for instance, also can be used to sustain cooperation in this infinitely repeated game and in others.

C. Games of Unknown Length

In addition to considering games of finite or infinite length, we can incorporate a more sophisticated tool to deal with games of unknown length. It is possible that, in some repeated games, players might not know for certain exactly how long their interaction will continue. They may, however, have some idea of the *probability* that the game will continue for another period. For example, our restaurants might believe that their repeated competition will continue only as long as their customers find *prix fixe* menus to be the dining-out experience of choice; if there were some probability each month that *à la carte* dinners would take over that role, then the nature of the game is altered.

Recall that the present value of a loss next month is already worth only $\delta = 1/(1 + r)$ times the amount earned. If in addition there is only a probability p (less than 1) that the relationship will actually continue to the next month, then next month's loss is worth only p times δ times the amount lost. For Xavier's Tapas, this means that the PV of the 36 lost with continued defecting is worth $36 \times \delta$ [the same as $36/(1 + r)$] when the game is assumed to be continuing with certainty but is worth only $36 \times p \times \delta$ when the game is assumed to be continuing with probability p. Incorporating the probability that the game may end next period means that the present value of the lost 36 is smaller, because $p < 1$, than it is when the game is definitely expected to continue (when p is *assumed* to equal 1).

The effect of incorporating p is that we now effectively discount future payoffs by the factor $p \times \delta$ instead of simply by δ. We call this **effective rate of return** R, where $1/(1 + R) = p \times \delta$, and R depends on p and δ as shown:[3]

$$1/(1 + R) = p\delta$$
$$1 = p\delta(1 + R)$$
$$R = \frac{1 - p\delta}{p\delta}.$$

[3]We could also express R in terms of r and p, in which case $R = (1 + r)/p - 1$.

With a 5% actual rate of return on investments (r = 0.05, and so $\delta = 1/1.05$ = 0.95) and a 50% chance that the game continues for an additional month (p = 0.5), then $R = [1 - (0.5)(0.95)]/(0.5)(0.95) = 1.1$, or 110%.

Now the high rates of return required to destroy cooperation (encourage defection) in these examples seem more realistic if we interpret them as effective rather than actual rates of return. It becomes conceivable that defecting forever, or even once, might actually be to one's benefit if there is a large enough probability that the game will end in the near future. Consider Xavier's decision whether to defect forever against a TFT-playing rival. Our earlier calculations showed that permanent defecting is beneficial only when r exceeds 1, or 100%. If Xavier's faces the 5% actual rate of return and the 50% chance that the game will continue for an additional month, as we assumed in the preceding paragraph, then the effective rate of return of 110% will exceed the critical value needed for it to continue defecting. Thus the cooperative behavior sustained by the TFT strategy can break down if there is a sufficiently large chance that the repeated game might be over by the end of the next period of play—that is, by a sufficiently small value of p.

D. General Theory

We can easily generalize the ideas about when it is worthwhile to defect against TFT-playing rivals so that you can apply them to any prisoners' dilemma game that you encounter. To do so, we use a table with general payoffs (delineated in appropriately measurable units) that satisfy the standard structure of payoffs in the dilemma as in Figure 11.3. The payoffs in the table must satisfy the relation H > C > D > L for the game to be a prisoners' dilemma, where C is the *cooperative* outcome, D is the payoff when both players *defect* from cooperation, H is the *high* payoff that goes to the defector when one player defects while the other cooperates, and L is the *low* payoff that goes to the loser (the cooperator) in the same situation.

In this general version of the prisoners' dilemma, a player's one-time gain from defecting is $(H - C)$. The single-period loss for being punished while you return to cooperation is $(C - L)$, and the per-period loss for perpetual defecting is $(C - D)$. To be as general as possible, we will allow for situations in which there is a probability $p < 1$ that the game continues beyond the next period and so we will discount payoffs using an effective rate of return of R per period. If $p = 1$, as would be the case when the game is guaranteed to continue, then $R = r$, the simple interest rate used in our preceding calculations. Replacing r with R, we find that the results attained earlier generalize almost immediately.

		COLUMN	
		Defect	Cooperate
ROW	Defect	D, D	H, L
	Cooperate	L, H	C, C

FIGURE 11.3 General Version of the Prisoners' Dilemma

We found earlier that a player defects exactly once against a rival playing TFT if the one-time gain from defecting $(H - C)$ exceeds the present value of the single-period loss from being punished (the PV of $C - L$). In this general game, that means that a player defects once against a TFT-playing opponent only if $(H - C) > (C - L)/(1 + R)$, or $(1 + R)(H - C) > C - L$, or

$$R > \frac{C - L}{H - C} - 1.$$

Similarly, we found that a player defects forever against a rival playing TFT only if the one-time gain from defecting exceeds the present value of the infinite sum of the per-period losses from perpetual defecting (where the per-period loss is $C - D$). For the general game, then, a player defects forever against a TFT-playing opponent only if $(H - C) > (C - D)/R$, or

$$R > \frac{C - D}{H - C}.$$

gain from defection (H - C)
loss for punishment { C - L, C - D

The three critical elements in a player's decision to defect, as seen in these two expressions, are the immediate gain from defection $(H - C)$, the future losses from punishment ($C - L$ or $C - D$ per period of punishment), and the value of the effective rate of return (R, which measures the importance of the present relative to the future). Under what conditions on these various values do players find it attractive to defect from cooperation?

First, assume that the values of the gains and losses from defecting are fixed. Then changes in R determine whether a player defects, and defection is more likely when R is large. Large values of R are associated with small values of p and small values of δ (and large values of r), so defection is more likely when the probability of continuation is low or the discount factor is low (or the interest rate is high). Another way to think about it is that defection is more likely when the future is less important than the present or when there is little future to consider; that is, defection is more likely when players are impatient or when they expect the game to end quickly.

Second, consider the case in which the effective rate of return is fixed, as is the one-period gain from defecting. Then changes in the per-period losses associated with punishment determine whether defecting is worthwhile. Here it is smaller values of $C - L$ or $C - D$ that encourage defection. In this case, defection is more likely when punishment is not very severe.[4]

Finally, assume that the effective rate of return and the per-period losses associated with punishment are held constant. Now players are more likely to defect when the gains, $H - C$, are high. This situation is more likely when defecting garners a player large and immediate benefits.

This discussion also highlights the importance of the detection of defecting. Decisions about whether to continue along a cooperative path depend on how long defecting might be able to go on before it is detected, on how accurately it is detected, and on how long any punishment can be made to last before an attempt is made to revert back to cooperation. Although our model does not incorporate these considerations explicitly, if defecting can be detected accurately and quickly, its benefit will not last long, and the subsequent cost will have to be paid more surely. Therefore the success of any trigger strategy in resolving a repeated prisoners' dilemma depends on how well (both in speed and accuracy) players can detect defecting. This is one reason that the TFT strategy is often considered dangerous; slight errors in the execution of actions or in the perception of those actions can send players into continuous rounds of punishment from which they may not be able to escape for a long time, until a slight error of the opposite kind occurs.

You can use all of these ideas to guide you in when to expect more cooperative behavior between rivals and when to expect more defecting and cutthroat actions. If times are bad and an entire industry is on the verge of collapse, for example, so that businesses feel that there is no future, competition may become fiercer (less cooperative behavior may be observed) than in normal times. Even if times are temporarily good but are not expected to last, firms may want to make a quick profit while they can, so cooperative behavior might again break down. Similarly, in an industry that emerges temporarily because of a quirk of fashion and is expected to collapse when fashion changes, we should expect less cooperation. Thus a particular beach resort might become the place to go, but all the hotels there will know that such a situation cannot last, and so they cannot afford to collude on pricing. If, on the other hand, the shifts in fashion are among products made by an unchanging group of companies in long-term relationships with each other, cooperation might persist. For example, even if all the children want cuddly bears one year and Power Ranger

[4] The costs associated with defection may also be smaller if information transmission is not perfect, as might be the case if there are many players, and so difficulties might arise in identifying the defector and in coordinating a punishment scheme. Similarly, gains from defection may be larger if rivals cannot identify a defection immediately.

action figures the next, collusion in pricing may occur if the same small group of manufacturers makes both items.

In Chapter 12, we will look in more detail at prisoners' dilemmas that arise in games with many players. We examine when and how players can overcome such dilemmas and achieve outcomes better for them all.

3 SOLUTIONS II: PENALTIES AND REWARDS

Although repetition is the major vehicle for the solution of the prisoners' dilemma, there are also several others that can be used to achieve this purpose. One of the simplest ways to avert the prisoners' dilemma in the one-shot version of the game is to inflict some direct **penalty** on the players when they defect. When the payoffs have been altered to incorporate the cost of the penalty, players may find that the dilemma has been resolved.[5]

Consider the husband–wife dilemma from Section 1. If only one player defects, the game's outcome entails one year in jail for the defector and 25 years for the cooperator. The defector, though, getting out of jail early, might find the cooperator's friends waiting outside the jail. The physical harm caused by those friends might be equivalent to an additional 20 years in jail. If so, and if the players account for the possibility of this harm, then the payoff structure of the original game has changed.

The "new" game, with the physical penalty included in the payoffs, is illustrated in Figure 11.4. With the additional 20 years in jail added to each player's sentence when one player confesses while the other denies, the game is completely different.

A search for dominant strategies in Figure 11.4 shows that there are none. A cell-by-cell check then shows that there are now two pure-strategy Nash equilibria. One of them is the (Confess, Confess) outcome; the other is the (Deny,

		WIFE	
		Confess	Deny
HUSBAND	Confess	10 yr, 10 yr	21 yr, 25 yr
	Deny	25 yr, 21 yr	3 yr, 3 yr

FIGURE 11.4 Prisoners' Dilemma with Penalty for the Lone Defector

[5] Note that we get the same type of outcome in the repeated-game case considered in Section 2.

Deny) outcome. Now each player finds that it is in his or her best interest to cooperate if the other is going to do so. The game has changed from being a prisoners' dilemma to an assurance game, which we studied in Chapter 4. Solving the new game requires selecting an equilibrium from the two that exist. One of them—the cooperative outcome—is clearly better than the other from the perspective of both players. Therefore it may be easy to sustain it as a focal point if some convergence of expectations can be achieved.

Notice that the penalty in this scenario is inflicted on a defector only when his or her rival does *not* defect. However, stricter penalties can be incorporated into the prisoners' dilemma, such as penalties for *any* confession. Such discipline typically must be imposed by a third party with some power over the two players, rather than by the other player's friends, because the friends would have little authority to penalize the first player when their associate also defects. If both prisoners are members of a special organization (such as a gang or a crime mafia) and the organization has a standing rule of never confessing to the police under penalty of extreme physical harm, the game changes again to the one illustrated in Figure 11.5.

Now the equivalent of an additional 20 years in jail is added to *all* payoffs associated with the Confess strategy. (Compare Figures 11.5 and 11.1.) In the new game, each player has a dominant strategy, as in the original game. The difference is that the change in the payoffs makes Deny the dominant strategy for each player. And (Deny, Deny) becomes the unique pure-strategy Nash equilibrium. The stricter penalty scheme achieved with third-party enforcement makes defecting so unattractive to players that the cooperative outcome becomes the new equilibrium of the game.

In larger prisoners' dilemma games, difficulties arise with the use of penalties. In particular, if there are many players and some uncertainty exists, penalty schemes may be more difficult to maintain. It becomes harder to decide whether actual defecting is taking place or it's just bad luck or a mistaken move. In addition, if there really is defecting, it is often difficult to determine the identity of the defector from among the larger group. And if the game is one shot, there is no opportunity in the future to correct a penalty that is too severe or to inflict a penalty once a defector has been identified. Thus penalties may be less

		WIFE	
		Confess	Deny
HUSBAND	Confess	30 yr, 30 yr	21 yr, 25 yr
	Deny	25 yr, 21 yr	3 yr, 3 yr

FIGURE 11.5 Prisoners' Dilemma with Penalty for Any Defecting

successful in large one-shot games than in the two-person game we consider here. We study prisoners' dilemmas with a large number of players in greater detail in Chapter 12.

A further interesting possibility arises when a prisoners' dilemma that has been solved with a penalty scheme is considered in the context of the larger society in which the game is played. It might be the case that, although the dilemma equilibrium outcome is bad for the players, it is actually good for the rest of society or for some subset of persons within the rest of society. If so, social or political pressures might arise to try to minimize the ability of players to break out of the dilemma. When third-party penalties are the solution to a prisoners' dilemma, as is the case with crime mafias that enforce a no-confession rule, for instance, society can come up with its own strategy to reduce the effectiveness of the penalty mechanism. The Federal Witness Protection Program is an example of a system that has been set up for just this purpose. The U.S. government removes the threat of penalty in return for confessions and testimony in court.

Similar situations can be seen in other prisoners' dilemmas, such as the pricing game between our two restaurants. The equilibrium there entailed both firms charging the low price of $20 even though they enjoy higher profits when charging the higher price of $26. Although the restaurants want to break out of this "bad" equilibrium—and we have already seen how the use of trigger strategies can help them do so—their customers are happier with the low price offered in the Nash equilibrium of the one-shot game. The customers then have an incentive to try to destroy the efficacy of any enforcement mechanism or solution process the restaurants might use. For example, because some firms facing prisoners' dilemma pricing games attempt to solve the dilemma through the use of a "meet the competition" or "price matching" campaign, customers might want to press for legislation banning such policies. We analyze the effects of such price-matching strategies in Section 7.B.

Just as a prisoners' dilemma can be resolved by penalizing defectors, it can also be resolved by rewarding cooperators. Because this solution is more difficult to implement in practice, we mention it only briefly.

The most important question is who is to pay the rewards. If it is a third party, that person or group must have sufficient interest of its own in the cooperation achieved by the prisoners to make it worth its while to pay out the rewards. A rare example of this occurred when the United States brokered the Camp David accords between Israel and Egypt by offering large promises of aid to both.

If the rewards are to be paid by the players themselves to each other, the trick is to make the rewards contingent (paid out only if the other player cooperates) and credible (guaranteed to be paid if the other player cooperates). Meeting these criteria requires an unusual arrangement; for example, the player making the promise should deposit the sum in advance in an escrow account held by an honorable and neutral third party, who will hand the sum over to the other player if she cooperates or return it to the promisor if the other defects. An end-of-chapter exercise shows how this type of arrangement can work.

4 SOLUTIONS III: LEADERSHIP

The third method of solution for the prisoners' dilemma pertains to situations in which one player takes on the role of leader in the interaction. In most examples of the prisoners' dilemma, the game is assumed to be symmetric. That is, all the players stand to lose (and gain) the same amount from defecting (and cooperation). However, in actual strategic situations, one player may be relatively "large" (a leader) and the other "small." If the size of the payoffs is unequal enough, so much of the harm from defecting may fall on the larger player that she acts cooperatively, even while knowing that the other will defect. Saudi Arabia, for example, played such a role as the "swing producer" in OPEC (Organization of Petroleum Exporting Countries) for many years; to keep oil prices high, it cut back on its output when one of the smaller producers, such as Libya, expanded.

As with the OPEC example, **leadership** tends to be observed more often in games between nations than in games between firms or individual persons. Thus our example for a game in which leadership may be used to solve the prisoners' dilemma is one played between countries. Imagine that the populations of two countries, Dorminica and Soporia, are threatened by a disease, Sudden Acute Narcoleptic Episodes (SANE). This disease strikes 1 person in every 2,000, or 0.05% of the population, and causes the victim to fall into a deep sleep state for a year.[6] There are no aftereffects of the disease, but the cost of a worker being removed from the economy for a year is \$32,000. Each country has a population of 100 million workers, so the expected number of cases in each is 50,000 ($0.0005 \times 100{,}000{,}000$), and the expected cost of the disease is \$1.6 billion to each ($50{,}000 \times 32{,}000$). The total expected cost of the disease worldwide—that is, in both Dorminica and Soporia—is then \$3.2 billion.

Scientists are confident that a crash research program costing \$2 billion will lead to a vaccine that is 100% effective. Comparing the cost of the research program with the worldwide cost of the disease shows that, from the perspective of the entire population, the research program is clearly worth pursuing. However, the government in each country must consider whether to fund the full research program on its own. They make this decision separately, but their

[6]Think of Rip Van Winkle or of Woody Allen in the movie *Sleeper*, but the duration is much shorter.

		SOPORIA	
		Research	No Research
DORMINICA	Research	−2, −2	−2, 0
	No Research	0, −2	−1.6, −1.6

FIGURE 11.6 Payoffs for Equal-Population SANE Research Game (\$billions)

decisions affect the outcomes for both countries. Specifically, if only one government chooses to fund the research, the population of the other country can access the information and use the vaccine without cost. But each government's payoff depends only on the costs incurred by its own population.

The payoff matrix for the noncooperative game between Dorminica and Soporia is shown in Figure 11.6. Each country chooses from two strategies, Research and No Research; payoffs show the costs to the countries, in billions of dollars, of the various strategy combinations. It is straightforward to verify that this game is a prisoners' dilemma and that each country has a dominant strategy to do no research.

But now suppose that the populations of the two countries are unequal, with 150 million in Dorminica and 50 million in Soporia. Then, if no research is funded by either government, the cost to Dorminica of SANE will be \$2.4 billion ($0.0005 \times 150{,}000{,}000 \times 32{,}000$) and the cost to Soporia will be \$0.8 billion ($0.0005 \times 50{,}000{,}000 \times 32{,}000$). The payoff matrix changes to the one illustrated in Figure 11.7.

In this version of the game, No Research is still the dominant strategy for Soporia. But Dorminica's best response is now Research. What has happened to change Dorminica's choice of strategy? Clearly, the answer lies in the unequal distribution of the population in this revised version of the game. Dorminica now stands to suffer such a large portion of the total cost of the disease that it finds it worthwhile to do the research on its own. This is true even though Dorminica knows full well that Soporia is going to be a free rider and get a share of the full benefit of the research.

		SOPORIA	
		Research	No Research
DORMINICA	Research	−2, −2	−2, 0
	No Research	0, −2	−2.4, −0.8

FIGURE 11.7 Payoffs for Unequal-Population SANE Research Game (\$billions)

The research game in Figure 11.7 is no longer a prisoners' dilemma. Here we see that the dilemma has, in a sense, been "solved" by the size asymmetry. The larger country chooses to take on a leadership role and provide the benefit for the whole world.

Situations of leadership in what would otherwise be prisoners' dilemma games are common in international diplomacy. The role of leader often falls naturally to the biggest or most well established of the players, a phenomenon labeled "the exploitation of the great by the small."[7] For many decades after World War II, for instance, the United States carried a disproportionate share of the expenditures of our defense alliances such as NATO and maintained a policy of relatively free international trade even when our partners, such as Japan and Europe, were much more protectionist. In such situations, it might be reasonable to suggest further that a large or well-established player may accept the role of leader because its own interests are closely tied to those of the players as a whole; if the large player makes up a substantial fraction of the whole group, such a convergence of interests would seem unmistakable. The large player would then be expected to act more cooperatively than might otherwise be the case.

5 SOLUTIONS IV: ASYMMETRIC INFORMATION

The final solution method we consider is one in which asymmetric information is introduced into a finitely repeated prisoners' dilemma. We saw in Section 2.A how an attempt to resolve the dilemma by repeated play would unravel by rollback reasoning if there were a fixed, finite number of plays. In actual play, however, even when players know exactly how long their interaction will last, they are able to sustain cooperation for quite a while; it unravels near the end when only a few rounds are left. When asked about their reasoning for cooperating in the early rounds, the players will usually say something such as, "I was willing to try and see if the other player was nice, and when this proved to be the case, I continued to cooperate until the time came to take advantage of the other's niceness." Of course the other player may not have been genuinely nice, but thinking along similar lines. As long as there is some chance that players in the dilemma are nice rather than selfish, it may pay even a selfish player to pretend to be nice. She can reap the higher payoffs from cooperation for a while and then also hope to exploit the gains from double crossing near the end of the sequence of plays. In this section, we will show how to explain such behavior more rigorously. If the above intuition suffices to satisfy your

[7]Mancur Olson, *The Logic of Collective Action* (Cambridge: Harvard University Press, 1965), p. 29.

curiosity about this solution, you can skip the rest of this section without loss of continuity.

A. General Expropriation Game

Note that this will be a game of asymmetric information. Players are of two types, selfish and nice. Each player knows his own type but not the type of the other player. Each is trying to infer the other's type from his actions. We solved such a game in Chapter 9, Section 5, where Fordor tried to infer Tudor's cost type from its choice of price. The same methods of analysis will work here, although the situation we have described above involves both players simultaneously trying to infer the other's type. Because the analysis of such a situation would get quite complicated, we will explain the ideas in a somewhat simpler example, in which only one player has the choice between being selfish and being nice. This type of game is sometimes called a one-person dilemma,[8] and is sometimes called a game of holdup or opportunism.[9]

Let us consider a specific situation in which a firm is deciding whether to invest in an emerging economy. The investment will entail an up-front cost of \$1 billion, and will then yield an operating profit of \$2 billion. It will also create spillover benefits to the country where the investment is located (the "host" country) of \$500 million. We will show all monetary amounts in billions, so these payoff numbers will be −1, 2, and 0.5, respectively.

After the investment is made, the host country's government will be tempted to change the rules so that it can collect the whole profit of 2 (billion) in addition to the spillover benefit of 0.5 (billion). That is, it can leave things as they are, accepting its payoff of 0.5, or it can expropriate the full profits from the firm's investment, thereby gaining itself a payoff of 2.5. The game tree in Figure 11.8 shows the host country's choices as *E* (for expropriate) and *NE* (for not expropriate); it has the opportunity to make this choice only after the firm has chosen to invest (*I*) rather than not to invest (*NI*).

A host country could achieve the expropriation outcome by nationalizing the local operation without compensating the foreign investor. Such expropriation of foreign investment has occurred quite often in history but is relatively rare these days. More common are indirect and partial expropriations that use changes in tax rules, limits on repatriation of profits, and so on. To keep matters simple we assume here that the expropriation of profit from the investing firm

[8]The most notable use of this terminology occurs in the works of Avner Greif; see his book *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade* (New York: Cambridge University Press, 2006).

[9] These concepts were developed and used by Oliver Williamson; see his book *The Economic Institutions of Capitalism* (New York: Free Press, 1987).

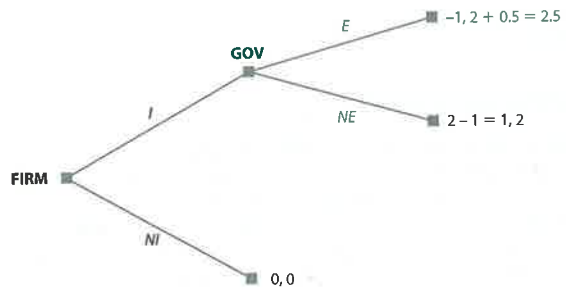

FIGURE 11.8 Simple Expropriation Game

by the host country is total. When applying the theory in other contexts, you will have to change the details to fit the specific situation.

In a single play of the game, then, rollback analysis of the game tree in Figure 11.8 shows that the host government will expropriate any available profits if the firm chooses to invest. Anticipating this choice, the firm will therefore choose not to invest. Similarly, when both players have full information about the other's possible and actual choices, the rollback equilibrium of the finitely repeated version of this game will entail no investment in any period; firms will not invest because they expect all profits to be expropriated. Just as in the finitely repeated prisoners' dilemma of Section 2.A, there will be no cooperation in equilibrium.

But what if we introduce an information asymmetry into this game? Specifically, assume that (host) governments come in two types, Opportunistic and Honorable, or O type and H type for short. Unable to distinguish the government's type, the firm must make its decision about whether to invest without knowing if the government is an O type or an H type. The former type of government will expropriate whenever that choice yields it a higher expected payoff than not expropriating; the latter type will never expropriate. Letting p denote the probability of the government being Honorable, we show the tree for the asymmetric information version of the expropriation game in Figure 11.9. There, in a single play of the game, the O type government will expropriate and the H type will not; therefore the firm's expected payoff from investing will be $2 \times p + 0 \times (1 - p) - 1 = 2p - 1$. The firm will invest if this expected payoff exceeds the expected payoff from not investing, 0; the firm will invest if $p > 1/2$.

Next, suppose the game is played repeatedly, but a fixed finite number of times and with no discounting across periods. The same host government will play in all periods, and its type will not change from one period to the next. Each period, a *new* firm gets the opportunity to make an investment. It observes whether firms invested in previous periods, and if so, whether the government

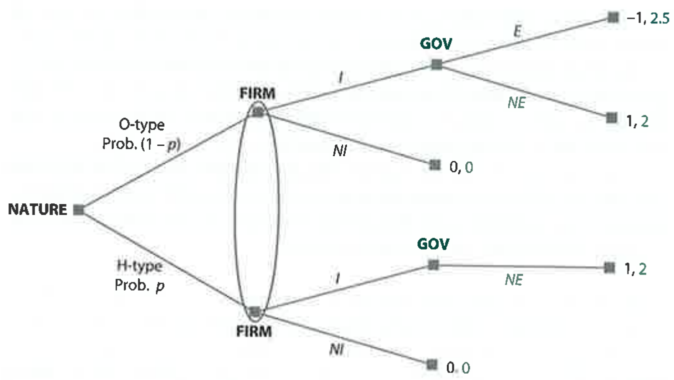

FIGURE 11.9 Expropriation Game with Asymmetric Information

expropriated. The prior probability held by the firm in the very first play of the game is that the government is H type with probability p.

In the repeated asymmetric information game, we will look for the following properties in the equilibrium:

1. Each new firm calculates an updated probability of the government being H type, using the previous period firm's prior belief along with the observed actions in that period and applying Bayes' rule. Its choice of whether to invest or not (*I* or *NI* for short) is optimal, given this updated probability.

2. The O-type government's decision whether to expropriate or not (*E* or *NE* for short) is optimal at all nodes in all periods, with the government recognizing the effect this choice will have on the probability calculations and actions of firms in future periods. An equilibrium that satisfies this properties will be a perfect Bayesian equilibrium (PBE) as defined in Chapter 9, Section 5.

B. Twice-Repeated Game with Asymmetric Information

Begin by considering the asymmetric information expropriation game when it is repeated for just two periods in total. To avoid confusion with our analysis of repeated games in Section 2.B, we use alphabetic, rather than numeric, labels here. The last period in actual time is labeled period Z; period Y is the one before that, and so on.

The firm's prior probability of the government being H type when entering the first period of play (period Y) is $p_Y = p$. Write p_Z for the prior probability of facing an H-type government when going into period Z (the second and final period of play). At this point, p_Z is just our notation for this probability; we will have to solve for its actual value as part of finding the equilibrium for the game.

We know already that in the period-Z game, the equilibrium actions for the government are to play *E* if it is O type and to play *NE* if it is H type. The firm plays *I* if $p_Z > 1/2$, or *NI* if $p_Z < 1/2$. It is indifferent between the two actions, and therefore willing to randomize between them if $p = 1/2$.

Now consider the period-Y game, which is the one played first in actual time. The possibilities for equilibrium play in that period will depend on the underlying value, p. We therefore distinguish three cases (the second of which will further subdivide) and consider each separately.

I. CASE I: $p > 1/2$ As in all situations, the H-type government will play *NE*. Given $p > 1/2$, the period-Y firm (firm Y) would play *I* even if the O-type government was playing *E*. Thus, it has a dominant strategy to play *I*. (It does best to play *I* against the H-type government and against the O type, regardless of the choice made by that government.) The O-type government has two possible strategies however. One, in which the O type plays *E*, would lead to separation, whereas the other, in which it plays *NE*, would result in pooling. We consider each possibility individually.

[1] Separation: We know that an H-type government plays *NE*, and firm Y plays *I*. Suppose an O-type government plays *E* in period *Y*; then the government's action would reveal its type. Can this set of strategies generate a separating equilibrium?

Given the strategies described, the firm investing in the second period (firm Z) will see that firm Y had been expropriated. Firm Z would then conclude that the government was O type for sure and would update the probability of it being H type to $p_Z = 0$. Therefore firm Z would not invest, so the O-type government would get 2.5 in period *Y* and 0 in period *Z*.

But what if the O-type government were to deviate and play *NE* in period *Y* instead? Observing *NE* in period *Y*, firm Z would update the probability of the government being H type to $p_Z = 1$. (Remember that in Nash equilibrium the firm will take the governments' equilibrium strategies as given, so it will believe that a government playing *NE* must be type H.) Therefore firm Z would play *I*, at which point the O-type government could play *E*. The government would then get 0.5 in period *Y* and 2.5 in period *Z*. This total payoff of 3 is better then the total of 2.5 the government gets by using its specified strategy, *E*, so that original strategy cannot be optimal. (Remember that we are not discounting across periods.) So we cannot get a separating equilibrium in the case of $p > 1/2$.

[2] Pooling: The second possibility is that the H-type government plays *NE*, firm Y plays *I*, and the O-type government also plays *NE*. Then the O type's action in period *Y* is the same as the H type's. Can these strategies constitute a pooling equilibrium?

Because both types of governments take the same action in period Y, no new information regarding type emerges for use by firm Z. Its Bayesian updating will lead to $p_Z = p_Y = p$. So firm Z will play *I* (because $p > 1/2$), and the O-type government will play *E*. Thus the O type gets 0.5 in period *Y* and 2.5 in period *Z* for a total of 3 when it follows the stipulated strategy.

If the O type were to deviate from its stated strategy and play *E* in period *Y*, this raises a question of how firm Z would update. The most natural assumption is that a choice of *E* would be interpreted as a sure indicator of O type, since *E* is not even a strategy available to the H type. Then firm Z, observing *E* in period *Y*, would play *NI*. The O-type government would get 2.5 in period *Y* and 0 in period *Z*. This total payoff is worse than what the O type gets from playing *NE* as stipulated, so the deviation is unprofitable. Pooling in the period-*Y* game is a perfect Bayesian equilibrium.

In this two-play case, the first play (period *Y*) has a different outcome from the single-play version of the game. Investment takes place, and profits are *not* expropriated by either type of government. (With just one play and with $p > 1/2$, investment would take place, but the type-O government would expropriate it.) Many observers would regard this outcome of the twice-played game as better than the single-play game, because actions are honorable, even though the sum of the players' payoffs is the same in both versions.

II. CASE II: $p < 1/2$ Again, the H-type government always plays *NE*. But with $p < 1/2$, it is no longer the case that firm Y has a dominant strategy to play *I*. Nor does it have a dominant strategy to play *NI*; if the O-type government pools and plays *NE*, firm Y's strategy *NI* could be part of a Nash equilibrium in period *Y*. Thus, we will have to consider all four possible combinations of pure strategies for firm Y (playing either *I* or *NI*) and the O-type government (playing either *E* or *NE*) to see which set or sets can be equilibria.

[1] Separation with investment: Consider first the set of strategies in which the H-type government plays *NE*, the O type plays *E*, and firm Y plays *I*. Although these strategies would result in a separation of types for the governments, they cannot be an equilibrium. Firm Y would have negative expected profit (because $p < 1/2$) and it would deviate to *NI*.

[2] Separation without investment: Now suppose that the H-type government plays *NE*, the O-type government plays *E*, and firm Y plays *NI*. Again, this would lead to separation if the set of strategies constitutes an equilibrium.

In this situation, firm Z gains no information about government type from the actions in period *Y* because no investment occurs. Thus, firm Z will also play

NI. Then the O type government's choices are irrelevant in both periods *Y* and *Z*, and it is indifferent between *E* and *NE*.

This makes the specified strategies a Nash equilibrium in period *Y*, but we do not have a perfect Bayesian equilibrium of the two-period game. Consider the off-equilibrium node where firm Y has played *I*. If the O-type government plays its stipulated equilibrium action *E*, that choice will reveal its type to firm Z, which will then play *NI*. So the O type's payoff would be 2.5 in period *Y*, and 0 in period *Z*. If instead the O type deviates to play *NE*, firm Z, which takes equilibrium strategies as given, will believe that the government is H type; that is, it will update to $p_Z = 1$. Therefore firm Z will invest. The O-type government can then expropriate and get a payoff of 2.5 in period *Z* to add to its payoff of 0.5 in period *Y*. The total payoff of 3 exceeds the payoff of 2.5 from playing *E*. This deviation is profitable to the O type, and even though this is true only when firm Y also deviates, it means that we cannot have a *perfect* Bayesian equilibrium of this type.

[3] Pooling with investment: Here we consider the possibility that both types of governments play *NE* while firm Y plays *I*. If these strategies are an equilibrium, we would have pooling of the two types of governments.

Given the stipulated strategies,the O-type government gets 0.5 in period *Y*. Because both governments play *NE*, there is no new information revealed for firm Z. Its updating leaves it with $p_Z = p_Y = p < 1/2$. Then firm Z plays *NI*, and the government gets 0 in period *Z*. The total payoff to the O-type government is 0.5 over the two periods.

If the O-type government were to deviate to *E* in period *Y*, it would get a payoff of 2.5 in that period. Its type would be revealed to firm Z, however, which would update to $p_Z = 0$ and therefore play *NI*. The government would get 0 in period *Z* and a total of 2.5 over the two periods. The O-type government's deviation is then profitable, and the originally stated strategies cannot be a Nash equilibrium.

[4] Pooling without investment: Our last possible set of pure strategies entails both types of governments playing *NE* and firm Y playing *NI*. These strategies cannot be an equilibrium, however, because firm Y would benefit by switching to *I*.

This analysis of the case of $p < 1/2$ shows that none of the four combinations of pure strategies for firm Y and the O-type government generate an equilibrium. With all of these pure-strategy combinations ruled out, we have to consider an equilibrium that entails mixing. With mixed strategies, we may be able to generate a *semiseparating equilibrium*.

[5] Semiseparation (with investment): Here we consider a possible equilibrium in which firm Y plays *I* while the O-type government mixes in period *Y*, playing *NE* with probability q_Y and *E* with probability $(1 - q_Y)$. In period *Z*, firm Z mixes, playing *I* with probability r_Z and *NI* with probability $(1 - r_Z)$, while the government

		GOVERNMENT ACTION		Sum of row
		NE	E	
GOV. TYPE	H	p	0	$p_Y = p$
	O	$(1-p)q_Y$	$(1-p)(1-q_Y)$	$1-p$
Sum of column		$p+(1-p)q_Y$	$(1-p)(1-q_Y)$	

FIGURE 11.10 Applying Bayes' Rule to the Expropriation Game

reverts to its true type; the O-type government plays E if the firm has invested. The values for q_Y and r_Z will be determined as part of our analysis of the equilibrium conditions. These conditions are the standard "opponent's-indifference" conditions; each player's mixture must keep the other indifferent between its pure actions.

In order for there to be an equilibrium with mixing by firm Z, the O-type government's period-*Y* mixture must keep firm Z indifferent between *I* and *NI*. For that, firm Z's Bayesian updating must yield $p_Z = 1/2$. What does this mean for the O type's choice of q_Y? To answer this question, we need to consider the probability table of types and actions illustrated in Figure 11.10. This table is similar to the ones we created in Chapter 9 when we explained Bayes' theorem in the Appendix and in the bluffing game of Section 5. Note that in the table the probability of observing an O-type government playing *NE* just equals the probability that the government is O type $(1 - p)$ times the probability that the O type chooses *NE* in its mixture (q_Y). The probability of observing an O type playing *E* is calculated similarly.

We can now use the table to determine how firm Z will update its probability that the government is H type. If firm Y's investment meets the government response *NE*, then Bayes' theorem states that the posterior probability of the government being H type (that is, firm Z's updated prior) will equal the probability of observing an H type playing *NE* divided by the sum of the probabilities associated with observing *NE*. The posterior probability is then $p/[p + (1 - p)q_Y]$.

Recall that to ensure mixing by firm Z, we need its updated probability that the government is H type to equal 1/2. Therefore we need

$$\frac{p}{p + (1-p)q_Y} = \frac{1}{2} \quad \text{or} \quad 2p = p + (1-p)q_Y \quad \text{or} \quad \frac{p}{1-p} = q_Y.$$

(Note that $p < 1/2$ ensures that $q_Y < 1$ will hold.) This condition specifies the appropriate level of q_Y for the O-type government's period-*Y* mixing.

Now we need to determine the correct mixture for firm Z. Its mix must keep the O-type government indifferent between *E* and *NE* in period *Y* (and therefore willing to mix in period Y). If the O-type government chooses E in period Y, it earns of payoff of 2.5 in that period but reveals its type. That revelation leads firm Z to play NI and the government gets a payoff of 0 in period Z, for a total payoff of 2.5. If the O type plays NE in period Y, it gets 0.5 in that period, and then firm Z will mix in period Z. With firm Z's mixing, the O-type government will get a payoff of 2.5 in period Z with probability r_Z (the probability that firm Z plays I) for a total payoff across the two periods of $0.5 + 2.5\ r_Z$. To keep the O-type government indifferent between E and NE in period Y, firm Z will want to choose the r_Z that equates these two payoffs. So firm Z needs $2.5 = 0.5 + 2.5\ r_Z$, or $r_Z = 0.8$.

We now have calculated equilibrium values for both q_Y and r_Z, but all of our analysis assumed that firm Y would choose I. If firm Y chose NI, there would be no action for the period-Y government to take and nothing to reveal its type even probabilistically. But we do need to verify that this assumption is valid.

To do so, we must consider firm Y's expected profit from investing in period Y. We know that firm Y's investment will not be expropriated if it meets an H-type government (probability p) or an O-type government choosing NE (probability $(1 - p)q_Y = (1 - p) \times p/(1 - p) = p$, using the solution above for q_Y). The total probability of meeting a government that will play NE is then $2p$. So firm Y gets expected profits of $(2p \times 2) - 1 = 4p - 1$. This expected profit is positive when $4p - 1 > 0$ or when $p > 1/4$. Therefore firm Y will invest if $p > 1/4$, and there will be a semiseparating equilibrium with the mixture probabilities calculated above. Note that the condition $p > 1/4$ is weaker than the $p > 1/2$ that was required to induce investment in the single-play version of this game. Thus repetition, even just two periods, increases the possibility of the good or cooperative outcome.

III. CASE III: $p = 1/2$ In this final case, firm Y will be indifferent between investing and not investing. This case is exceptional, being just on the borderline between the case of $p > 1/2$ (where we found a pooling equilibrium in which firm Y invests and neither type of government expropriates) and the case $1/2 > p > 1/4$ (where we found a semiseparating equilibrium in which firm Y invests and the O-type government randomizes between E and NE). As p rises to 1/2 in the range of the semiseparating equilibria, the probability of the O-type government choosing NE, $q_Y = p/(1 - p)$, rises to 1. So the two cases on either side of $p = 1/2$ converge to the same outcome. Therefore we will regard the case $p = 1/2$ as a limiting case of the first two and we will not go into its details separately.

C. Thrice-Repeated Game

Our analysis in Section 5.B showed that going to a twice-repeated version of the expropriation game increased the likelihood that the cooperative outcome would be observed in equilibrium. We now consider additional repetitions, starting specifically with the case of a three-period game. Counting backward again, the first period of play will be labeled period X, with periods Y and Z being the second and third (or final) periods, as before. Here we can show that the equilibrium has the following features:

Case a: If $p > 1/4$, there will be a pooling equilibrium in period X where firm X invests and even an opportunistic government does not expropriate.

Case b: If $1/4 > p > 1/8$, there will be a semiseparating equilibrium where firm X (the first to play) invests, the O-type government randomizes in its response, and firm Y also randomizes.

Note that the range of values of p where investment takes place and is not expropriated in the first period of play has expanded geometrically (in powers of 1/2) with the increase in the number of repetitions. This pattern would continue if we were to add more repetitions of the game.

The details of the analysis verifying the equilibrium strategies are similar to those of the twice-repeated case, so we omit most of them. But we want to emphasize and check two key issues in the thrice-repeated case.

First, we need to verify the optimality of nonexpropriation in Case a; it must be optimal for the O-type government to play NE in period X (the first period of play) when the initial probability is $p_X = p > 1/4$. If the O-type government does play NE, it will get 0.5 in period X (remember, firm X plays I). This action pools it with the H-type government, so it reveals no new information about type to firm Y. The game in period Y therefore has the same $p_Y = p > 1/4$. Our analysis in Section 5.B above showed that the O-type government's total payoff over periods Y and Z is 2.5 when $1/2 > p_Y > 1/4$ and 3 when $p_Y > 1/2$. Therefore, over the three periods the O-type government gets 3 when $1/2 > p_Y > 1/4$ and 3.5 when $p_Y > 1/2$. If it deviated and chose E in the very first play (period X), it would get 2.5 in that period, but it would reveal its type to firms Y and Z and so get a payoff of 0 thereafter. The deviation from NE in period X is therefore not profitable, and pooling in period X is an equilibrium in this case.

Second, we must check the condition from Case b that guarantees that randomization is sustained when $1/4 > p$. The O-type government's period-X (first play) randomization should keep firm Y indifferent about investing. By the analysis for the twice-repeated case, this indifference will be ensured when firm Y's Bayesian updating yields $p_Y = 1/4$. Therefore, as above, we need $p/[p+(1 - p)q_X] = 1/4$. (The equilibrium entry probability, r_Y, in firm Y's mixture is similarly calculated to be 0.8.)

Finally, firm X will indeed invest if its expected profit is positive. Firm X's profits are not expropriated with probability p (that it meets an H-type government) plus $(1 - p)q_X$ (that it meets an O-type government playing NE). Then firm X's expected profit is $[p + (1 - p)q_X] \times 2 - 1 = 8p - 1$, where we have made use of the condition defining q_X that was derived in the preceding paragraph. This expected profit is positive, and firm X does invest, if $p > 1/8$. This condition on investment in the first period of play is even weaker than that found in the two-stage game.

Further repetitions will follow the same pattern. If the game is played N times where N is large, there will be a pooling equilibrium with investment and no expropriation in the initial $(N - n)$ periods, where n is defined as the smallest integer that makes $p < (1/2)^n$ true. In the following $(n - 1)$ periods there will be semiseparating equilibria. The O-type government's randomization in one of these following periods may yield expropriation, in which case later firms will not invest. Otherwise, in the last period (period Z) the firm will be indifferent between investing and not investing because its updated p_Z will exactly equal 1/2. (This result follows from the observation of *NE* in the period-Y semiseparating equilibrium.) But in period Z, an O-type government will play *NE* for sure.

In an exercise at the end of this chapter, we will guide you through a more general formulation of this game, with the payoffs and probabilities denoted by algebraic symbols instead of specific numbers, to show that the idea underlying this solution is perfectly general. The corresponding two-sided dilemma game is harder to solve, and we merely refer ambitious readers to the original article.[10]

6 EXPERIMENTAL EVIDENCE

Numerous people have conducted experiments in which subjects compete in prisoners' dilemma games against each other.[11] Such experiments show that cooperation can and does occur in such games, even in repeated versions of known and finite length. Many players start off by cooperating and continue to cooperate for quite a while, as long as the rival player reciprocates. Only in the last few plays of a finite game does defecting seem to creep in. Although this behavior goes against the reasoning of rollback, it can be "profitable" if sustained for a reasonable length of time. The pairs get higher payoffs than would rational, calculating strategists who defect from the very beginning.

[10] David Kreps, Paul Milgrom, John Roberts, and Robert Wilson, "Rational Cooperation in a Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory,* vol. 27 (1982), pp. 245–252.

[11] The literature on experiments involving the prisoners' dilemma game is vast. A brief overview is given by Alvin Roth in *The Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995), pp. 26–28. Journals in both psychology and economics can be consulted for additional references. For some examples of the outcomes that we describe, see Kenneth Terhune, "Motives, Situation, and Interpersonal Conflict Within Prisoners' Dilemmas," *Journal of Personality and Social Psychology Monograph Supplement,* vol. 8, no. 30 (1968), pp. 1–24; and R. Selten and R. Stoecker, "End Behavior in Sequences of Finite Prisoners' Dilemma Supergames," *Journal of Economic Behavior and Organization,* vol. 7 (1986), pp. 47–70. Robert Axelrod's *Evolution of Cooperation* (New York: Basic Books, 1984) presents the results of his computer-simulation tournament for the best strategy in an infinitely repeated dilemma.

Such observed behavior can be rationalized in different ways. Perhaps the players are not sure that the relationship will actually end at the stated time. Perhaps they believe that their reputations for cooperation will carry over to other similar games against the same opponent or other opponents. Perhaps they think it possible that their opponents are naive cooperators, and they are willing to risk a little loss in testing this hypothesis for a couple of plays. If successful, the experiment will lead to higher payoffs for a sufficiently long time.

In some laboratory experiments, players engage in multiple-round games, each round consisting of a given finite number of repetitions. All of the repetitions in any one round are played against the same rival, but each new round is played against a new opponent. Thus there is an opportunity to develop cooperation with an opponent in each round and to "learn" from preceding rounds when devising one's strategy against new opponents as the rounds continue. These situations have shown that cooperation lasts longer in early rounds than in later rounds. This result suggests that the theoretical argument on the unraveling of cooperation, based on the use of rollback, is being learned from experience of the play itself over time as players begin to understand the benefits and costs of their actions more fully. Another possibility is that players learn simply that they want to be the first to defect, and so the timing of the initial defection occurs earlier as the number of rounds played increases.

Suppose you were playing a game with a prisoners' dilemma structure and found yourself in a cooperative mode with the known end of the relationship approaching. When should you decide to defect? You do not want to do so too early, while a lot of potential future gains remain. But you also do not want to leave it until too late in the game, because then your opponent might preempt you and leave you with a low payoff for the period in which she defects. In fact, your decision about when to defect cannot be deterministic. If it were, your opponent would figure it out and defect in the period before you planned to do so. If no deterministic choice is feasible, then the unraveling of cooperation must include some uncertainty, such as mixed strategies, for both players. Many thrillers whose plots hinge on tenuous cooperation among criminals or between informants and police acquire their suspense precisely because of this uncertainty.

Examples of the collapse of cooperation as players near the end of a repeated game are observed in numerous situations in the real world, as well as in the laboratory. The story of a long-distance bicycle (or foot) race is one such example. There may be a lot of cooperation for most of the race, as players take turns leading and letting others ride in their slipstreams; nevertheless, as the finish line looms, each participant will want to make a dash for the tape. Similarly, signs saying "no checks accepted" often appear in stores in college towns each spring near the end of the semester.

Computer-simulation experiments have matched a range of very simple to very complex contingent strategies against each other in two-player prisoners' dilemmas. The most famous of them were conducted by Robert Axelrod at the University of Michigan. He invited people to submit computer programs that specified a strategy for playing a prisoners' dilemma repeated a finite but large number (200) of times. There were 14 entrants. Axelrod held a "league tournament" that pitted pairs of these programs against one another, in each case for a run of the 200 repetitions. The point scores for each pairing and its 200 repetitions were kept, and each program's scores over all its runs against different opponents were added up to see which program did best in the aggregate against all other programs. Axelrod was initially surprised when "nice" programs did well; none of the top eight programs were ever the first to defect. The winning strategy turned out to be the simplest program: Tit-for-tat, submitted by the Canadian game theorist Anatole Rapoport. Programs that were eager to defect in any particular run got the defecting payoff early but then suffered repetitions of mutual defections and poor payoffs. On the other hand, programs that were always nice and cooperative were badly exploited by their opponents. Axelrod explains the success of Tit-for-tat in terms of four properties: it is at once forgiving, nice, provocable, and clear.

In Axelrod's words, one does well in a repeated prisoners' dilemma to abide by these four simple rules: "Don't be envious. Don't be the first to defect. Reciprocate both cooperation and defection. Don't be too clever."[12] Tit-for-tat embodies each of the four ideals for a good, repeated prisoners' dilemma strategy. It is not envious; it does not continually strive to do better than the opponent, only to do well for itself. In addition, Tit-for-tat clearly fulfills the admonitions not to be the first to defect and to reciprocate, defecting only in retaliation to the opponent's preceding defection and always reciprocating in kind. Finally, Tit-for-tat does not suffer from being overly clever; it is simple and understandable to the opponent. In fact, it won the tournament not because it helped players achieve high payoffs in any individual game—the contest was not about "winner takes all"—but because it was always close; it simultaneously encourages cooperation and avoids exploitation, whereas other strategies cannot.

Axelrod then announced the results of his tournament and invited submissions for a second round. Here, people had a clear opportunity to design programs that would beat Tit-for-tat. The result: Tit-for-tat won again! The programs that were cleverly designed to beat it could not beat it by very much, and they did poorly against one another. Axelrod also arranged a tournament

[12]Axelrod, *Evolution of Cooperation*, p. 110.

of a different kind. Instead of a league where each program met each other program once, he ran a game with a whole population of programs, with a number of copies of each program. Each type of program met an opponent randomly chosen from the population. Those programs that did well were given a larger proportion of the population; those that did poorly had their proportion in the population reduced. This was a game of evolution and natural selection, which we will study in greater detail in Chapter 13. But the idea is simple in this context, and the results are fascinating. At first, nasty programs did well at the expense of nice ones. But as the population became nastier and nastier, each nasty program met other nasty programs more and more often, and they began to do poorly and fall in numbers. Then Tit-for-tat started to do well and eventually triumphed.

However, Tit-for-tat has some flaws. Most importantly, it assumes no errors in execution of the strategy. If there is some risk that the player intends to play the cooperative action but plays the defecting action in error, then this action can initiate a sequence of retaliatory defecting actions that locks two Tit-for-tat programs playing one another into a bad outcome; another error is required to rescue them from this sequence. When Axelrod ran a third variant of his tournament, which provided for such random mistakes, Tit-for-tat could be beaten by even "nicer" programs that tolerated an occasional episode of defecting to see if it was a mistake or a consistent attempt to exploit them and retaliated only when convinced that it was not a mistake.[13]

Interestingly, a twentieth-anniversary competition modeled after Axelrod's original contest and run in 2004 and 2005 generated a new winning strategy.[14] Actually, the winner was a set of strategies designed to recognize one another during play so that one would become docile in the face of the other's continued defections. (The authors likened their approach to a situation in which prisoners manage to communicate with each other by tapping on their cell walls.) This collusion meant that some of the strategies submitted by the winning team did very poorly, whereas others did spectacularly well, a testament to the value of working together. Of course Axelrod's contest did not permit multiple submissions, so such strategy sets were ineligible, but the winners of the recent competition argue that with no way to preclude coordination, strategies such as those they submitted should have been able to win the original competition as well.

[13]For a description and analysis of Axelrod's computer simulations from the biological perspective, see Matt Ridley, *The Origins of Virtue* (New York: Penguin Books, 1997), pp. 61, 75. For a discussion of the difference between computer simulations and experiments using human players, see John K. Kagel and Alvin E. Roth, *Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995), p. 29.

[14]See Wendy M. Grossman, "New Tack Wins Prisoner's Dilemma," *Wired*, October 13, 2004. Available at http://www.wired.com/culture/lifestyle/news/2004/10/65317 (accessed 6/14/08).

7 REAL-WORLD DILEMMAS

Games with the prisoners' dilemma structure arise in a surprisingly varied number of contexts in the world. Although we would be foolish to try to show you every possible instance in which the dilemma can arise, we take the opportunity in this section to consider in detail three specific examples from a variety of fields of study. One example comes from evolutionary biology, a field that we will study in greater detail in Chapter 13. A second example describes the policy of "price matching" as a solution to a prisoners' dilemma pricing game. And a final example concerns international environmental policy and the potential for repeated interactions to mitigate the prisoners' dilemma in this situation.

A. Evolutionary Biology

In our first example, we consider a game known as the bowerbirds' dilemma, from the field of evolutionary biology.[15] Male bowerbirds attract females by building intricate nesting spots called bowers, and female bowerbirds are known to be particularly choosy about the bowers built by their prospective mates. For this reason, male bowerbirds often go out on search-and-destroy missions aimed at ruining other males' bowers. While they are out, however, they run the risk of losing their own bower to the beak of another male. The ensuing competition between male bowerbirds and their ultimate choice regarding whether to maraud or guard has the structure of a prisoners' dilemma game.

Ornithologists have constructed a table that shows the payoffs in a two-bird game with two possible strategies, Maraud and Guard. That payoff table is shown in Figure 11.11. GG represents the benefits associated with Guarding when the rival bird also Guards; GM represents the payoff from Guarding when the rival bird is a Marauder. Similarly, MM represents the benefits associated with Marauding when the rival bird also is a Marauder; MG represents the payoff

		BIRD 2	
		Maraud	Guard
BIRD 1	Maraud	MM, MM	MG, GM
	Guard	GM, MG	GG, GG

FIGURE 11.11 Bowerbirds' Dilemma

[15] Larry Conik, "Science Classics: The Bowerbird's Dilemma," *Discover,* October 1994.

from Marauding when the rival bird Guards. Careful scientific study of bowerbird matings led to the discovery that MG > GG > MM > GM. In other words, the payoffs in the bowerbird game have exactly the same structure as the prisoners' dilemma. The birds' dominant strategy is to maraud, but when both choose that strategy, they end up in equilibrium each worse off than if they had both chosen to guard.

In reality, the strategy used by any particular bowerbird is not actually the result of a process of rational choice on the part of the bird. Rather, in evolutionary games, strategies are assumed to be genetically "hardwired" into individual organisms, and payoffs represent reproductive success for the different types. Then equilibria in such games define the type of population that naturalists can expect to observe—all marauders, for instance, if Maraud is a dominant strategy as in Figure 11.11. This equilibrium outcome is not the best one, however, given the existence of the dilemma. In constructing a solution to the bowerbirds' dilemma, we can appeal to the repetitive nature of the interaction in the game. In the case of the bowerbirds, repeated play against the same or different opponents in the course of several breeding seasons can allow you, the bird, to choose a flexible strategy based on your opponent's last move. Contingent strategies such as tit-for-tat can be, and often are, adopted in evolutionary games to solve exactly this type of dilemma. We will return to the idea of evolutionary games and provide detailed discussions of their structure and equilibrium outcomes in Chapter 13.

B. Price Matching

Now we return to a pricing game, in which we consider two specific stores engaged in price competition with each other, using identical price-matching policies. The stores in question, Toys "R" Us and Kmart, are both national chains that regularly advertise prices for name-brand toys (and other items). In addition, each store maintains a published policy that guarantees customers that it will match the advertised price of any competitor on a specific item (model and item numbers must be identical) as long as the customer provides the competitor's printed advertisement.[16]

For the purposes of this example, we assume that the firms have only two possible prices that they can charge for a particular toy (Low or High). In addition, we use hypothetical profit numbers and further simplify the analysis by

[16] The price-matching policy at Toys "R" Us is printed and posted prominently in all stores. A simple phone call confirmed that Kmart has an identical policy. Similar policies are appearing in many industries, including that for credit cards where "interest rate matching" has been observed. See Aaron S. Edlin, "Do Guaranteed-Low-Price Policies Guarantee High Prices, and Can Antitrust Rise to the Challenge?" *Harvard Law Review,* vol. 111, no. 2 (December 1997), pp. 529–575.

		KMART	
		Low	High
TOYS "R" US	Low	2,000, 2,000	4,000, 0
	High	0, 4,000	3,000, 3,000

FIGURE 11.12 Toys "R" Us and Kmart Toy Pricing

assuming that Toys "R" Us and Kmart are the only two competitors in the toy market in a particular city—Billings, Montana, for example.

Suppose, then, that the basic structure of the game between the two firms can be illustrated as in Figure 11.12. If both firms advertise low prices, they split the available customer demand and each earns $2,000. If both advertise high prices, they split a market with lower sales, but their markups end up being large enough to let them each earn $3,000. Finally, if they advertise different prices, then the one advertising a high price gets no customers and earns nothing, whereas the one advertising a low price earns $4,000.

The game illustrated in Figure 11.12 is clearly a prisoners' dilemma. Advertising and selling at a low price is the dominant strategy for each firm, although both would be better off if each advertised and sold at the high price. But as mentioned earlier, each firm actually makes use of a third pricing strategy: a price-matching guarantee to its customers. How does the inclusion of such a policy alter the prisoners' dilemma that would otherwise exist between these two firms?

Consider the effects of allowing firms to choose among pricing low, pricing high, and price matching. The Match strategy entails advertising a high price but promising to match any lower advertised price by a competitor; a firm using Match then benefits from advertising high if the rival firm does so also, but it does not suffer any harm from advertising a high price if the rival advertises a low price. We can see this in the payoff structure for the new game, shown in Figure 11.13. In that table, we see that a combination of one firm playing Low while the other plays Match is equivalent to both playing Low, while a combination of one firm playing High while the other plays Match (or both playing Match) is equivalent to both playing High.

Using our standard tools for analyzing simultaneous-play games shows that High is weakly dominated by Match for both players and that once High is eliminated, Low is weakly dominated by Match also. The resulting Nash equilibrium entails both firms using the Match strategy. In equilibrium, both firms earn $3,000—the profit level associated with both firms pricing high in the original game. The addition of the Match strategy has allowed the firms to emerge from the prisoners' dilemma that they faced when they had only the choice between two simple pricing strategies, Low or High.

		KMART		
		Low	High	Match
TOYS "R" US	Low	2,000, 2,000	4,000, 0	2,000, 2,000
	High	0, 4,000	3,000, 3,000	3,000, 3,000
	Match	2,000, 2,000	3,000, 3,000	3,000, 3,000

FIGURE 11.13 Toys "R" Us and Kmart Toy Pricing

How did this happen? The Match strategy acts as a penalty mechanism. By guaranteeing to match Kmart's low price, Toys "R" Us substantially reduces the benefit that Kmart achieves by advertising a low price while Toys "R" Us is advertising a high price. In addition, promising to meet Kmart's low price hurts Toys "R" Us, too, because the latter has to accept the lower profit associated with the low price. Thus the price-matching guarantee is a method of penalizing both players whenever either one defects. This is just like the crime mafia example discussed in Section 3, except that this penalty scheme—and the higher equilibrium prices that it supports—is observed in markets in virtually every city in the country.

Actual empirical evidence of the detrimental effects of these policies is available but limited, and some research has found evidence of lower prices in markets with such policies.[17] However, more recent experimental evidence does support the collusive effect of price-matching policies. This result should put all customers on alert.[18] Even though stores that match prices promote their policies in the name of competition, the ultimate outcome when all firms use such policies can be better for the firms than if there were no price matching at all, and so customers can be the ones who are hurt.

C. International Environmental Policy: The Kyoto Protocol

Our final example pertains to the international climate control agreement known as the Kyoto Protocol. Negotiated by the United Nations Framework Convention

[17] J. D. Hess and Eitan Gerstner present evidence of increased prices as a result of price-matching policies in "Price-Matching Policies: An Empirical Case," *Managerial and Decision Economics*, vol. 12 (1991), pp. 305–315. Contrary evidence is provided by Arbatskaya, Hviid, and Shaffer, who find that the effect of matching policies is to lower prices; see Maria Arbatskaya, Morten Hviid, and Greg Shaffer, "Promises to Match or Beat the Competition: Evidence from retail Tire Prices," *Advances in Applied Microeconomics*, vol. 8: Oligopoly (New York: JAI Press, 1999), pp. 123–138.

[18] See Subhasish Dugar, "Price-Matching Guarantees and Equilibrium Selection in a Homogeneous Product Market: An Experimental Study," *Review of Industrial Organization*, vol. 30 (2007), pp. 107–119.

		THEM	
		Cut Emissions	Don't Cut
US	Cut Emissions	−1, −1	−20, 0
	Don't Cut	0, −20	−12, −12

FIGURE 11.14 Greenhouse Gas Emissions Game

on Climate Change in 1997 as a tool for reducing greenhouse gas emissions, it went in to effect in 2005 and is due to expire in 2012. Over 170 countries have signed on to the treaty, although the United States is noticeably absent from the list. Ongoing meetings continue to work on a plan for extending the protocol beyond its current end date.

The difficulty in achieving global reduction in greenhouse gas emissions comes in part from the prisoners' dilemma nature of the interaction. Any individual country will have no incentive to reduce its own emissions, knowing that if it does so alone it bears significant costs with little benefit to overall climate change. If others do reduce their emissions the first country cannot be stopped from enjoyiong the benefits of the others' actions.

Consider the emissions reduction problem as a game played between two countries, Us and Them. Estimates generated by the British government's Office on Climate Change suggest that coordinated action may come at a cost of about 1% of GDP per nation, whereas coordinated inaction could cost each nation between 5 and 20% of GDP, perhaps 12% on average.[19] By extension, the cost to cutting emissions on your own may be at the high end of the inaction estimate (20%), but holding back and letting the other country cut emissions could entail virtually no cost to you at all. We can then summarize the situation between Us and Them using the game table in Figure 11.14, where payoffs represent changes in GDP for each country.

The game in Figure 11.14 is indeed a prisoners' dilemma. Both countries have a dominant strategy to refuse to cut their emissions. The single Nash equilibrium occurs when neither country cuts emissions, but they suffer as a group as a result of the ensuing climate change. From this analysis we should expect little or no progress in greenhouse gas emissions reduction.

This interpretation of the problem inherent in the Kyoto Protocol has been challenged by recent research from Michael Liebriech, who argues that the game

[19]See Nicholas Stern, *The Economics of Climate Change: The Stern Review* (Cambridge: Cambridge University Press, 2007).

is not a one-off interaction and that countries repeatedly interact and negotiate additional amendments to the existing agreement.[20] He argues that the iterated nature of this game makes it amenable to solution by way of contingent strategies and that countries should use strategies that embody the four critical properties of TFT as outlined by Axelrod and described in Section 6 above. Specifically, countries are encouraged to employ strategies that are "nice" (signing on to the protocol and beginning emissions reductions), "retaliatory" (employing mechanisms to punish those that do not do their part), "forgiving" (welcoming to those newly accepting the protocol), and "clear" (specifying actions and reactions).

Liebriech assesses the actions of current players, including the European Union, the United States, and developing countries (as a group), and provides some suggestions for improvements. He explains that the European Union does well with nice, forgiving, and clear but not with retaliation, so other countries will do best to defect when interacting with the European Union. One solution would be for the European Union to institute carbon-related import taxes or another retaliatory-type policy for dealing with recalcitrant trade partners. The United States, on the other hand, ranks high on retaliatory and forgiving, given its history of such behavior following the end of the cold war. But it has not been nice or clear, at least on the national level (individual states may behave differently), giving other countries an incentive to retaliate against it quickly and painfully, if possible. The solution is for the United States to make a meaningful commitment to carbon-emission reduction, a standard conclusion in most policy circles. Developing countries are described as not nice (negotiating no carbon limits for themselves), retaliatory, unclear, and quite unforgiving. A more beneficial strategy, argues Liebriech, would be for these countries—particularly China, India, and Brazil—to make clear their commitment to sharing in international efforts to affect climate change; this approach would leave them less subject to retaliation and more likely to benefit from a global improvement in climatic outlook.

The general conclusion is that the process of international carbon emissions reduction does fit the profile of a prisoners' dilemma game. But the future of global greenhouse gas emissions should not be considered a lost cause simply because of the prisoners' dilemma aspects of the one-time interaction. Repeated play among the nations involved in the Kyoto Protocol negotiations make the game amenable to solutions by way of contingent (nice, clear, and forgiving, but also retaliatory) strategies.

[20]Michael Liebriech presents his analysis of the Kyoto Protocol as an iterated prisoners' dilemma in his paper "How to Save the Planet: Be Nice, Retaliatory, Forgiving and Clear," New Energy Finance White Paper, September 11, 2007. Available at www.newenergyfinance.com/docs/Press/NEF-WP_Carbon-Game-Theory_05.pdf (accessed 9/11/08).

SUMMARY

The prisoners' dilemma is probably the most famous game of strategy. Each player has a dominant strategy (to Defect), but the equilibrium outcome is worse for all players than when each uses her dominated strategy (to Cooperate). The best-known solution to the dilemma is *repetition of play*. In a finitely played game, the *present value* of future cooperation is eventually zero, and rollback yields an equilibrium with no cooperative behavior. With infinite play (or an uncertain end date), cooperation can be achieved with the use of an appropriate contingent strategy such as *tit-for-tat (TFT)* or the *grim strategy*; in either case, cooperation is possible only if the present value of cooperation exceeds the present value of defecting. More generally, the prospects of "no tomorrow" or of short-term relationships lead to decreased cooperation among players.

The dilemma can also be "solved" with *penalty* schemes that alter the payoffs for players who defect from cooperation when their rivals are cooperating or when others also are defecting. A third solution method arises if a large or strong player's loss from defecting is greater than the available gain from cooperative behavior on that player's part. Allowing for asymmetric information in the dilemma can lead to some cooperation, even in finitely repeated games.

Experimental evidence suggests that players often cooperate longer than theory might predict. Such behavior can be explained by incomplete knowledge of the game on the part of the players or by their views regarding the benefits of cooperation. Tit-for-tat has been observed to be a simple, nice, provocable, and forgiving strategy that performs very well on the average in repeated prisoners' dilemmas.

Prisoners' dilemmas arise in a variety of contexts. Specific examples from international environmental policy, evolutionary biology, and product pricing show how to explain and predict actual behavior by using the framework of the prisoners' dilemma.

KEY TERMS

compound interest (404)
contingent strategy (401)
discount factor (404)
effective rate of return (405)
grim strategy (401)
infinite horizon (404)
leadership (412)
penalty (409)
present value (PV) (403)
punishment (401)
repeated play (399)
tit-for-tat (TFT) (401)
trigger strategy (401)

SOLVED EXERCISES

S1. "If a prisoners' dilemma is repeated 100 times, and both players know how many repetitions to expect, they are sure to achieve their cooperative outcome." True or false? Explain and give an example of a game that illustrates your answer.

S2. Consider a two-player game between Child's Play and Kid's Korner, each of which produces and sells wooden swing sets for children. Each player can set either a high or a low price for a standard two-swing, one-slide set. If they both set a high price, each receives profits of \$64,000 per year. If one sets a low price and the other sets a high price, the low-price firm earns profits of \$72,000 per year, while the high-price firm earns \$20,000. If they both set a low price, each receives profits of \$57,000.

(a) Verify that this game has a prisoners' dilemma structure by looking at the ranking of payoffs associated with the different strategy combinations (both cooperate, both defect, one defects, and so on). What are the Nash-equilibrium strategies and payoffs in the simultaneous-play game if the players meet and make price decisions only once?

(b) If the two firms decide to play this game for a fixed number of periods—say, for 4 years—what would each firm's total profits be at the end of the game? (Don't discount.) Explain how you arrived at your answer.

(c) Suppose that the two firms play this game repeatedly forever. Let each of them use a grim strategy in which they both price high unless one of them "defects," in which case they price low for the rest of the game. What is the one-time gain from defecting against an opponent playing such a strategy? How much does each firm lose, in each future period, after it defects once? If $r = 0.25$ ($\delta = 0.8$), will it be worthwhile for them to cooperate? Find the range of values of r (or δ) for which this strategy is able to sustain cooperation between the two firms.

(d) Suppose the firms play this game repeatedly year after year, neither expecting any change in their interaction. If the world were to end after 4 years, without either firm having anticipated this event, what would each firm's total profits (not discounted) be at the end of the game? Compare your answer here with the answer in part (b). Explain why the two answers are different, if they are different, or why they are the same, if they are the same.

(e) Suppose now that the firms know that there is a 10% probability that one of them may go bankrupt in any given year. If bankruptcy occurs, the repeated game between the two firms ends. Will this knowledge change the firms' actions when $r = 0.25$? What if the probability of a bankruptcy increases to 35% in any year?

S3. A firm has two divisions, each of which has its own manager. Managers of these divisions are paid according to their effort in promoting productivity in their divisions. The payment scheme is based on a comparison of the two outcomes. If both managers have expended "high effort," each earns \$150,000 a year. If both have expended "low effort," each earns "only" \$100,000 a year. But if one of the two managers shows "high effort" whereas the other shows "low effort," the "high effort" manager is paid \$150,000 plus a \$50,000 bonus, but the second ("low effort") manager gets a reduced salary (for subpar performance in comparison with her competition) of \$80,000. Managers make their effort decisions independently and without knowledge of the other manager's choice.

(a) Assume that expending effort is costless to the managers and draw the payoff table for this game. Find the Nash equilibrium of the game and explain whether the game is a prisoners' dilemma.

(b) Now suppose that expending high effort is costly to the managers (such as a costly signal of quality). In particular, suppose that "high effort" costs an equivalent of \$60,000 a year to a manager who chooses this effort level. Draw the game table for this new version of the game and find the Nash equilibrium. Explain whether the game is a prisoners' dilemma and how it has changed from the game in part (a).

(c) If the cost of high effort is equivalent to \$80,000/year, how does the game change from that described in part (b)? What is the new equilibrium? Explain whether the game is a prisoners' dilemma and how it has changed from the games in parts (a) and (b).

S4. You have to decide whether to invest \$100 in a friend's enterprise, where in a year's time the money will increase to \$130. You have agreed that your friend will then repay you \$120, keeping \$10 for himself. But instead he may choose to run away with the whole \$130. Any of your money that you don't invest in your friend's venture you can invest elsewhere safely at the prevailing rate of interest r, and get $\$100(1 + r)$ next year.

(a) Draw the game tree for this situation and show the rollback equilibrium.

Next, suppose this game is played repeatedly infinitely often. That is, each year you have the opportunity to invest another \$100 in your friend's enterprise, and the agreement is to split the resulting \$130 in the manner already described. From the second year onward, you get to make your decision of whether to invest with your friend in the light of whether he made the agreed repayment the preceding year. The rate of interest between any two successive periods is r, the same as the outside rate of interest and the same for you and your friend.

(b) For what values of r can there be an equilibrium outcome of the repeated game, in which each period you invest with your friend and he repays as agreed?

(c) If the rate of interest is 10% per year, can there be an alternative profit-splitting agreement that is an equilibrium outcome of the infinitely repeated game, where each period you invest with your friend and he repays as agreed?

S5. Recall the example from Exercise S3 in which two division managers' choices of High or Low effort levels determine their salary payments. In part (b) of that exercise, the cost of exerting High effort is assumed to be \$60,000 a year. Suppose now that the two managers play the game in part (b) of Exercise S3 repeatedly for many years. Such repetition allows scope for an unusual type of cooperation in which one is designated to choose High effort while the other chooses Low. This cooperative agreement requires that the High-effort manager make a side payment to the Low-effort manager so that their payoffs are identical.

(a) What size side payment guarantees that the final payoffs of the two managers are identical? How much does each manager earn in a year in which the cooperative agreement is in place?

(b) Cooperation in this repeated game entails each manager's choosing her assigned effort level and the High-effort manager making the designated side payment. Defection entails refusing to make the side payment. Under what values of the rate of return can this agreement sustain cooperation in the managers' repeated game?

S6. Consider the game of chicken in Chapter 4, with slightly more general payoffs (Figure 4.14 had $k = 1$):

		DEAN	
		Swerve	Straight
JAMES	Swerve	0, 0	$-1, k$
	Straight	$k, -1$	$-2, -2$

Suppose this game is played repeatedly, every Saturday evening. If $k < 1$, the two players stand to benefit by cooperating to play (Swerve, Swerve) all the time, whereas if $k > 1$, they stand to benefit by cooperating so that one plays Swerve and the other plays Straight, taking turns to go Straight in alternate weeks. Can either type of cooperation be sustained?

S7. Recall the example from Exercise S8 of Chapter 5, where South Korea and Japan compete in the market for production of VLCCs. As in parts (a) and (b) of that exercise, the cost of building ships is \$30 (million) in each country, and the demand for ships is $P = 180 - Q$, where $Q = q_{Korea} + q_{Japan}$.

(a) Previously, we found the Nash equilibrium for the game. Now find the collusive outcome. What total quantity should be set by the two countries in order to maximize their joint profit?

(b) Suppose the two countries produce equal quantities of VLCCs, so that they earn equal shares of this collusive profit. How much profit would each country earn? Compare this profit with the amount they would earn in the Nash equilibrium.

(c) Now suppose the two countries are in a repeated relationship. Once per year, they choose production quantities, and each can observe the amount its rival produced in the previous year. They wish to cooperate to sustain the collusive profit levels found in part (b). In any one year, one of them can defect from the agreement. If one of them holds the quantity at the agreed level, what is the best defecting quantity for the other? What are the resulting profits?

(d) Write down a matrix that represents this game as a prisoners' dilemma.

(e) For what interest rates will collusion be sustainable when the two countries use grim (defect forever) strategies?

UNSOLVED EXERCISES

U1. Two people, Baker and Cutler, play a game in which they choose and divide a prize. Baker decides how large the total prize should be; she can choose either \$10 or \$100. Cutler chooses how to divide the prize chosen by Baker; Cutler can choose either an equal division or a split where she gets 90% and Baker gets 10%. Write down the payoff table of the game and find its equilibria for each of the following situations:

(a) When the moves are simultaneous.

(b) When Baker moves first.

(c) When Cutler moves first.

(d) Is this game a prisoners' dilemma? Why or why not?

U2. Consider a small town that has a population of dedicated pizza eaters but is able to accommodate only two pizza shops, Donna's Deep Dish and Pierce's Pizza Pies. Each seller has to choose a price for its pizza, but for simplicity, assume that only two prices are available: high and low. If a high price is set, the sellers can achieve a profit margin of \$12 per pie; the low price yields a profit margin of \$10 per pie. Each store has a loyal captive customer base that will buy 3,000 pies per week, no matter what price is charged by either store. There is also a floating demand of 4,000 pies per week. The people who buy these pies are price conscious and will go to the store with the lower price; if both stores charge the same price, this demand will be split equally between them.

(a) Draw the game table for the pizza-pricing game, using each store's profits per week (in thousands of dollars) as payoffs. Find the Nash equilibrium of this game and explain why it is a prisoners' dilemma.

(b) Now suppose that Donna's Deep Dish has a much larger loyal clientele that guarantees it the sale of 11,000 (rather than 3,000) pies a week. Profit margins and the size of the floating demand remain the same. Draw the payoff table for this new version of the game and find the Nash equilibrium.

(c) How does the existence of the larger loyal clientele for Donna's Deep Dish help "solve" the pizza stores' dilemma?

U3. A town council consists of three members who vote every year on their own salary increases. Two Yes votes are needed to pass the increase. Each member would like a higher salary but would like to vote against it herself because that looks good to the voters. Specifically, the payoffs of each are as follows:

Raise passes, own vote is No: 10
Raise fails, own vote is No: 5
Raise passes, own vote is Yes: 4
Raise fails, own vote is Yes: 0

Voting is simultaneous. Write down the (three-dimensional) payoff table, and show that in the Nash equilibrium the raise fails unanimously. Examine how a repeated relationship among the members can secure them salary increases every year if (i) every member serves a 3-year term, (ii) every year in rotation one of them is up for reelection, and (iii) the townspeople have short memories, remembering only the votes on the salary-increase motion of the current year and not those of past years.

U4. Consider the following game, which comes from James Andreoni and Hal Varian at the University of Michigan.[21] A neutral referee runs the game. There are two players, Row and Column. The referee gives two cards to each: 2 and 7 to Row and 4 and 8 to Column. This is common knowledge. Then, playing simultaneously and independently, each player is asked to hand over to the referee either his high card or his low card. The referee hands out payoffs—which come from a central kitty, not from the players' pockets—

[21] James Andreoni and Hal Varian, "Preplay Contacting in the Prisoners' Dilemma," *Proceedings of the National Academy of Sciences*, vol. 96, no. 19 (September 14, 1999), pp. 10933–10938.

that are measured in dollars and depend on the cards that he collects. If Row chooses his Low card, 2, then Row gets \$2; if he chooses his High card, 7, then Column gets \$7. If Column chooses his Low card, 4, then Column gets \$4; if he chooses his High card, 8, then Row gets \$8.

(a) Show that the complete payoff table is as follows:

		COLUMN	
		Low	High
ROW	Low	2, 4	10, 0
	High	0, 11	8, 7

(b) What is the Nash equilibrium? Verify that this game is a prisoners' dilemma.

Now suppose the game has the following stages. The referee hands out cards as before; who gets what cards is common knowledge. At stage I, each player, out of his own pocket, can hand over a sum of money, which the referee is to hold in an escrow account. This amount can be zero but cannot be negative. When both have made their Stage I choices, these are publicly disclosed. Then at stage II, the two make their choices of cards, again simultaneously and independently. The referee hands out payoffs from the central kitty in the same way as in the single-stage game before. In addition, he disposes of the escrow account as follows. If Column chooses his high card, the referee hands over to Column the sum that Row put into the account; if Column chooses his low card, Row's sum reverts back to him. The disposition of the sum that Column deposited depends similarly on Row's card choice. All these rules are common knowledge.

(c) Find the rollback (subgame-perfect) equilibrium of this two-stage game. Does it resolve the prisoners' dilemma? What is the role of the escrow account?

U5. Glassworks and Clearsmooth compete in the local market for windshield repairs. The market size (total available profits) is \$10 million per year. Each firm can choose whether to advertise on local television. If a firm chooses to advertise in a given year, it costs that firm \$3 million. If one firm advertises and the other doesn't, then the former captures the whole market. If both firms advertise, they split the market 50:50. If both firms choose not to advertise, they also split the market 50:50.

(a) Suppose the two windshield-repair firms know they will compete for just one year. Write down the payoff matrix for this game. Find the Nash equilibrium strategies.

(b) Suppose the firms play this game for five years in a row, and they know that at the end of five years, both firms plan to go out of business. What is the subgame-perfect equilibrium for this five-period game? Explain.

(c) What would be a "tit-for-tat" strategy in the game described in part (b)?

(d) Suppose the firms play this game repeatedly forever, and suppose that future profits are discounted with an interest rate of 20% per year. Can you find a subgame-perfect equilibrium that involves higher annual payoffs than the equilibrium in part (b)? If so, explain what strategies are involved. If not, explain why not.

U6. Consider the pizza stores introduced in Exercise U2, Donna's Deep Dish and Pierce's Pizza Pies. Suppose that they are not constrained to choose from only two possible prices, but that they can choose a specific value for price to maximize profits. Suppose further that it costs \$3 to make each pizza (for each store) and that experience or market surveys have shown that the relation between sales (Q) and price (P) for each firm is as follows:

$$Q_{\text{Pierce}} = 12 - P_{\text{Pierce}} + 0.5P_{\text{Donna}}.$$

Then profits per week (Y, in thousands of dollars) for each firm are:

$$Y_{\text{Pierce}} = (P_{\text{Pierce}} - 3)\ Q_{\text{Pierce}} = (P_{\text{Pierce}} - 3)\ (12 - P_{\text{Pierce}} + 0.5P_{\text{Donna}}),$$
$$Y_{\text{Donna}} = (P_{\text{Donna}} - 3)\ Q_{\text{Donna}} = (P_{\text{Donna}} - 3)\ (12 - P_{\text{Donna}} + 0.5\ P_{\text{Pierce}}).$$

(a) Use these profit functions to determine each firm's best-response rule, as in Chapter 5, and use the best-response rules to find the Nash equilibrium of this pricing game. What prices do the firms choose in equilibrium? How much profit per week does each firm earn?

(b) If the firms work together and choose a joint best price, P, then the profit of each will be:

$$Y_{\text{Donna}} = Y_{\text{Pierce}} = (P - 3)\ (12 - P + 0.5\ P) = (P - 3)\ (12 - 0.5\ P).$$

What price do they choose to maximize joint profits?

(c) Suppose the two stores are in a repeated relationship, trying to sustain the joint profit-maximizing prices calculated in part (b). They print new menus each month and thereby commit themselves to prices for the whole month. In any one month, one of them can defect from the agreement. If one of them holds the price at the agreed level, what is the best defecting price for the other? What are its resulting profits? For what interest rates will their collusion be sustainable by using grim-trigger strategies?

U7. Now we extend the analysis of Exercise S7 to allow for defecting in a collusive triopoly. Exercise S9 of Chapter 5 finds the Nash outcome of a VLCC triopoly of Korea, Japan, and China.

(a) Now find the collusive outcome of the triopoly. That is, what total quantity should be set by the three countries collectively in order to maximize their joint profit?

(b) Assume that under the collusive outcome found in part (a), the three countries produce equal quantities of VLCCs, so that each earns an equal share of the collusive profit. How much profit would each country earn? Compare this profit with the amount each earns in the Nash outcome.

(c) Now suppose the three countries are in a repeated relationship. Once per year, they choose production quantities, and each can observe the amount its rivals produced in the previous year. They wish to cooperate to sustain the collusive profit levels found in part (b). In any one year, one of them can defect from the agreement. If the other two countries are expected to produce their share of the collusive outcome found in parts (a) and (b), what is the best defecting quantity for the third to produce? What is the resulting profit for a defecting country when it produces the optimal defecting quantity while the other two produce their collusive quantities?

(d) Of course, the year after one country defects, both of its rivals will also defect. They will all find themselves back at the Nash outcome (permanently, if they use grim-trigger strategies). How much does the defecting country stand to gain in one year of defecting from the collusive outcome? How much will the defecting country then lose in every subsequent year from earning the Nash profit instead of the collusive profit?

(e) For what interest rates will collusion be sustainable if the three countries are using grim-trigger strategies? Is this set of interest rates larger or smaller than that found in the duopoly case discussed in Exercise S7, part (e)? Why?

Appendix: Infinite Sums

The computation of present values requires us to determine the current value of a sum of money that is paid to us in the future. As we saw in Section 2 of Chapter 11, the present value of a sum of money—say, x—that is paid to us n months from now is just $x/(1+r)^n$, where r is the appropriate monthly rate of return. But the present value of a sum of money that is paid to us next month and every following month in the foreseeable future is more complicated to determine. In that case, the payments continue infinitely, and so there is no defined end to the sum of present values that we need to compute. To compute the present value of this flow of payments requires some knowledge of the mathematics of the summation of infinite series.

Consider a player who stands to gain \$36 this month from defecting in a prisoners' dilemma but who will then lose \$36 every month in the future as a result of her choice to continue defecting while her opponent punishes her (using the tit-for-tat, or TFT, strategy). In the first of the future months—the first for which there is a loss and the first for which values need to be discounted—the present value of her loss is $36/(1+r)$; in the second future month, the present value of the loss is $36/(1+r)^2$; in the third future month, the present value of the loss is $36/(1+r)^3$. That is, in each of the n future months that she incurs a loss from defecting, that loss equals $36/(1+r)^n$.

We could write out the total present value of all of her future losses as a large sum with an infinite number of components,

$$PV = \frac{36}{1+r} + \frac{36}{(1+r)^2} + \frac{36}{(1+r)^3} + \frac{36}{(1+r)^4} + \frac{36}{(1+r)^5} + \frac{36}{(1+r)^6} + \cdots,$$

or we could use summation notation as a shorthand device and instead write

$$PV = \sum_{n=1}^{\infty} \frac{36}{(1+r)^n}.$$

This expression, which is equivalent to the preceding one, is read as "the sum, from n equals 1 to n equals infinity, of 36 over $(1+r)$ to the nth power." Because 36 is a common factor—it appears in each term of the sum—it can be pulled out to the front of the expression. Thus we can write the same present value as

$$PV = 36 \times \sum_{n=1}^{\infty} \frac{36}{(1+r)^n}.$$

We now need to determine the value of the sum within the present-value expression to calculate the actual present value. To do so, we will simplify our notation by switching to the *discount factor* δ in place of $1/(1+r)$. Then the sum that we are interested in evaluating is

$$\sum_{n=1}^{\infty} \delta^n.$$

It is important to note here that $\delta = 1/(1+r) < 1$ because r is strictly positive.

An expert on infinite sums would tell you, after inspecting this last sum, that it converges to the finite value $\delta/(1-\delta)$.[1] Convergence is guaranteed because increasingly large powers of a number less than 1, δ in this case, become smaller

[1]An infinite series *converges* if the sum of the values in the series approaches a specific value, getting closer and closer to that value as additional components of the series are included in the sum. The series *diverges* if the sum of the values in the series gets increasingly larger (more negative) with each addition to the sum. Convergence requires that the components of the series get progressively smaller.

and smaller, approaching zero as n approaches infinity. The later terms in our present value, then, decrease in size until they get sufficiently small that the series approaches (but technically never exactly reaches) the particular value of the sum. Although a good deal of more sophisticated mathematics is required to deduce that the convergent value of the sum is $\delta/(1 - \delta)$, proving that this is the correct answer is relatively straightforward.

We use a simple trick to prove our claim. Consider the sum of the first m terms of the series, and denote it by S_m. Thus

$$S_m = \sum_{n=1}^{\infty} \delta^n = \delta + \delta^2 + \delta^3 + \cdots + \delta^{m-1} + \delta^m.$$

Now we multiply this sum by $(1 - \delta)$ to get

$$\begin{aligned}(1-\delta)S_m &= \delta + \delta^2 + \delta^3 + \cdots + \delta^{m-1} + \delta^m \\ &\quad - \delta^2 - \delta^3 - \delta^4 - \cdots - \delta^m - \delta^{m-1} \\ &= \delta - \delta^{m-1}.\end{aligned}$$

Dividing both sides by $(1 - \delta)$, we have

$$S_m = \frac{\delta - \delta^{m+1}}{1 - \delta}.$$

Finally we take the limit of this sum as m approaches infinity to evaluate our original infinite sum. As m goes to infinity, the value of δ^{m+1} goes to zero because very large and increasing powers of a number less than 1 get increasingly small but stay nonnegative. Thus as m goes to infinity, the right-hand side of the preceding equation goes to $\delta/(1 - \delta)$, which is therefore the limit of S^m as m approaches infinity. This completes the proof.

We need only convert back into r to be able to use our answer in the calculation of present values in our prisoners' dilemma games. Because $\delta = 1/(1 + r)$, it follows that

$$\frac{\delta}{1-\delta} = \frac{1/(1+r)}{r/(1+r)} = \frac{1}{r}.$$

The present value of an infinite stream of \$36s earned each month, starting next month, is then

$$36 \times \sum_{n=1}^{\infty} \frac{1}{(1+r)^n} = \frac{36}{r}.$$

This is the value that we use to determine whether a player should defect forever in Section 2 of Chapter 11. Notice that incorporating a probability of continuation, $p \leq 1$, into the discounting calculations changes nothing in the summation procedure used here. We could easily substitute R for r in the preceding calculations, and $p\delta$ for the discount factor, δ.

Remember that you need to find present values only for losses (or gains) incurred (or accrued) *in the future*. The present value of \$36 lost today is just \$36. So if you wanted the present value of a stream of losses, all of them \$36, that begins *today*, you would take the \$36 lost today and add it to the present value of the stream of losses in the future. We have just calculated that present value as $36/r$. Thus the present value of the stream of lost \$36s, including the \$36 lost today, would be $36 + 36/r$, or $36[(r + 1)/r]$, which equals $36/(1 - \delta)$. Similarly, if you wanted to look at a player's stream of profits under a particular contingent strategy in a prisoners' dilemma, you would not discount the profit amount earned in the very first period; you would only discount those profit figures that represent money earned in future periods.