setter offers 6 in period 1, and that the payoffs are -24 for the agenda setter and -40 for the voter. Assume in this exercise that the voter chooses the higher acceptable policy when indifferent. If you are courageous, show that this policy is uniquely optimal when the agenda setter's discount factor is slightly less than 1 instead of 1.

(b) Suppose now that both players are rational. Show that the agenda setter's utility is higher and the voter's utility is lower than in question (a). What point does this comparison illustrate? (See Ingberman 1985 and Rosenthal 1990.)

References

Basu, K. 1988. Strategic irrationality in extensive games. *Mathematical Social Sciences* 15: 247–260.

Basu, K. 1990. On the non-existence of a rationality definition for extensive games. *International Journal of Game Theory*.

Binmore, K. 1987. Modeling rational players: Part I. *Economics and Philosophy* 3: 179–214.

Binmore, K. 1988. Modeling rational players: Part II. *Economics and Philosophy* 4: 9–55.

Bonanno, G. 1988. The logic of rational play in extensive games of perfect information. Mimeo, University of California, Davis.

Brander, J., and B. Spencer. 1985. Export subsidies and market share rivalry. *Journal of International Economics* 18: 83–100.

Dixit, A. 1979. A model of duopoly suggesting a theory of entry barriers. *Bell Journal of Economics* 10: 20–32.

Eaton, J., and G. Grossman. 1986. Optimal trade and industrial policy under oligopoly. *Quarterly Journal of Economics* 101: 383–406.

Eckel, C., and C. Holt. 1989. Strategic voting in agenda-controlled committee experiments. *American Economic Review* 79: 763–773.

Elmes, S., and P. Reny. 1988. The equivalence of games with perfect recall. Mimeo.

Fudenberg, D., and D. Kreps. 1988. A theory of learning, experimentation, and equilibrium in games. Mimeo, MIT.

Fudenberg, D., D. Kreps, and D. Levine. 1988. On the robustness of equilibrium refinements. *Journal of Economic Theory* 44: 354–380.

Fudenberg, D., and D. Levine. 1990. Steady-state learning and self-confirming equilibrium. Mimeo.

Harris, C. 1985. Existence and characterization of perfect equilibrium in games of perfect information. *Econometrica* 53: 613–627.

Hellwig, M., and W. Leininger. 1987. On the existence of subgame-perfect equilibrium in infinite-action games of perfect information. *Journal of Economic Theory*.

Helpman, E., and P. Krugman. 1989. *Trade Policy and Market Structure*. MIT Press.

Ingberman, D. 1985. Running against the status-quo. *Public Choice* 146: 14–44.

Kreps, D., and R. Wilson. 1982. Sequential equilibria. *Econometrica* 50: 863–894.

Kuhn, H. 1953. Extensive games and the problem of information. *Annals of Mathematics Studies*, no. 28. Princeton University Press.

Kydland, F., and E. Prescott. 1977. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85: 473–491.

Luce, R., and H. Raiffa. 1957. *Games and Decisions*. Wiley.

Mankiw, G. 1988. Recent developments in macroeconomics: A very quick refresher course. *Journal of Money, Credit, and Banking* 20: 436–459.

Rabin, M. 1988. Consistency and robustness criteria for game theory. Mimeo, MIT.

Reny, P. 1986. Rationality, common knowledge, and the theory of games. Ph.D. Dissertation, Princeton University.

Romer, T., and H. Rosenthal. 1978. Political resource allocation, controlled agendas, and the status-quo. *Public Choice* 33: 27–44.

Rosenthal, H. 1990. The setter model. In *Readings in the Spatial Theory of Elections*, ed. Enelow and Hinich. Cambridge University Press.

Rosenthal, R. 1981. Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory* 25: 92–100.

Schelling, T. 1960. *The Strategy of Conflict*. Harvard University Press.

Selten, R. 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft* 12: 301–324.

Selten, R. 1975. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4: 25–55.

Shepsle, K. 1981. Structure-induced equilibrium and legislative choice. *Public Choice* 37: 503–520.

Spence, A. M. 1977. Entry, capacity, investment and oligopolistic pricing. *Bell Journal of Economics* 8: 534–544.

Zermelo, E. 1913. Über eine Anwendung der Mengenlehre auf der Theorie des Schachspiels. In Proceedings of the Fifth International Congress on Mathematics.

4 Applications of Multi-Stage Games with Observed Actions

4.1 Introduction[†]

In chapter 3 we introduced a class of extensive-form games that we called "multi-stage games with observed actions," where the players move simultaneously within each stage and know the actions that were chosen in all past stages. Although these games are very special, they have been used in many applications in economics, political science, and biology. The repeated games we study in chapter 5 belong to this class, as do the games of resource extraction, preemptive investment, and strategic bequests discussed in chapter 13. This chapter develops a basic fact about dynamic optimization and presents a few interesting examples of multi-stage games. The chapter concludes with discussions of what is meant by "open-loop" and "closed-loop" equilibria, of the notion of iterated conditional dominance, and of the relationship between equilibria of finite-horizon and infinite-horizon games.

Recall that in a multi-stage game with observed actions the history h^t at the beginning of stage t is simply the sequence of actions $(a^0, a^1, \ldots, a^{t-1})$ chosen in previous periods, and that a pure strategy s_i for player i is a sequence of maps s_i^t from histories h^t to actions a_i^t in the feasible sets $A_i(h^t)$. Player i's payoff u_i is a function of the terminal history h^{T+1}, i.e., of the entire sequence of actions from the initial stage 0 through the terminal stage T, where T is sometimes taken to be infinite. In some of the examples of this section, payoffs take the special form of the discounted sum $\sum_{t=0}^{T} \delta_i^t g_i^t(a^t)$ of per-period payoffs $g_i^t(a^t)$.

Section 4.3 presents a first look at the subclass of repeated games, where the payoffs are given by averages as above and where the sets of feasible actions at each stage and the per-period payoffs are independent of previous play and time, so that the "physical environment" of the game is memoryless. Nevertheless, the fact that the game is repeated means that the players can condition their current play on the past play of their opponents, and indeed there can be equilibria in strategies of this kind. Section 4.3 considers only a few examples of repeated games, and does not try to characterize all the equilibria of the examples it examines; chapter 5 gives a more thorough treatment.

In this chapter we consider mostly games with an infinite horizon as opposed to a horizon that is long but finite. Games with a long but finite horizon represent a situation where the horizon is long but well foreseen; infinite-horizon games describe a situation where players are fairly uncertain as to which period will be the last. This latter assumption seems to be a better model of many situations with a large number of stages; we will say more about this point when discussing some of the examples.

When the horizon is infinite, the set of subgame-perfect equilibria cannot be determined by backward induction from the terminal date, as it can be in the finitely repeated prisoner's dilemma and in any finite game of perfect information. As we will see, however, subgame perfection does lead to very strong predictions in some infinite-horizon games with a great many Nash equilibria, such as the bargaining model of Rubinstein (1982) and Ståhl (1972). A key feature of this model and of some of the others we will discuss is that, although the horizon is not *a priori* bounded, there are some actions, such as accepting an offer or exiting from a market, that effectively "end the game." These games have been applied to the study of exit from a declining industry, noncooperative bargaining, and the introduction of new technology, among other topics. Section 4.4 discusses the Rubinstein-Ståhl alternating-offer bargaining game, where there are many ways the game can end, corresponding to the various possible agreements the players can reach. Section 4.5 discusses the class of simple timing games, where the players' only decision is when to stop the game and not the "way" to stop it. We have not tried to give a thorough survey of the many applications of games with absorbing states; our purpose is to introduce some of the flavor of the ideas involved.

Section 4.6 introduces the concept of "iterated conditional dominance," which extends the concept of backward induction to games with a potentially infinite number of stages. As we will see, the unique subgame-perfect equilibria of several of the examples we discuss in this chapter can be understood as the consequence of there being a single strategy profile that survives the weaker condition of iterated conditional dominance. Section 4.7 discusses the relationship between open-loop equilibria and closed-loop equilibria, which are the equilibria of two different information structures for the same "physical game." Section 4.8 discusses the relationship between the equilibria of finite- and infinite-horizon versions of the "same game."

The last two sections are more technical than the rest of the chapter, and could be skipped in a first course. Sections 4.3–4.6 are meant to be examples of the uses of the theory we developed in chapter 3. Most courses would want to cover at least one of these sections, but it is unnecessary to do all of them. Section 4.2, though, is used in chapters 5 and 13; it develops a fact that is very useful in determining whether a strategy profile is subgame perfect.

4.2 The Principle of Optimality and Subgame Perfection†

To verify that a strategy profile of a multi-stage game with observed actions is subgame perfect, it suffices to check whether there are any histories h^t where some player i can gain by deviating from the actions prescribed by s_i at h^t and conforming to s_i thereafter. Since this "one-stage-deviation principle" is essentially the principle of optimality of dynamic programming, which is based on backward induction, it helps illustrate how subgame perfection extends the idea of backward induction. We split the observation into two parts, corresponding to finite- and infinite-horizon games; some readers may prefer to read the first proof and take the second one on faith, although both are quite simple. For notational simplicity, we state the principle for pure strategies; the mixed-strategy counterpart is straightforward.

Theorem 4.1 (one-stage-deviation principle for finite-horizon games) In a finite multi-stage game with observed actions, strategy profile s is subgame perfect if and only if it satisfies the one-stage-deviation condition that no player i can gain by deviating from s in a single stage and conforming to s thereafter. More precisely, profile s is subgame perfect if and only if there is no player i and no strategy $\hat{s}_i$ that agrees with s_i except at a single t and h^t, and such that $\hat{s}_i$ is a better response to s_{-i} than s_i conditional on history h^t being reached.[1]

Proof The necessity of the one-stage-deviation condition ("only if") follows from the definition of subgame perfection. (Note that the one-stage-deviation condition is not necessary for *Nash* equilibrium, as a Nash-equilibrium profile may prescribe suboptimal responses at histories that do not occur when the profile is played.) To see that the one-stage-deviation condition is sufficient, suppose to the contrary that profile s satisfies the condition but is not subgame perfect. Then there is a stage t and a history h^t such that some player i has a strategy $\hat{s}_i$ that is a better response to s_{-i} than s_i is in the subgame starting at h^t. Let $\hat{t}$ be the largest t' such that, for some $h^{t'}$, $\hat{s}_i(h^{t'}) \neq s_i(h^{t'})$. The one-stage-deviation condition implies $\hat{t} > t$, and since the game is finite, $\hat{t}$ is finite as well. Now consider an alternative strategy $\tilde{s}_i$ that agrees with $\hat{s}_i$ at all $t < \hat{t}$ and follows s_i from stage $\hat{t}$ on. Since $\hat{s}_i$ agrees with s_i from $\hat{t} + 1$ on, the one-stage-deviation condition implies that $\tilde{s}_i$ is as good a response as $\hat{s}_i$ in every subgame starting at $\hat{t}$, so $\tilde{s}_i$ is as good a response as $\hat{s}_i$ in the subgame starting at t with history h^t. If $\hat{t} = t + 1$, then $\tilde{s}_i = s_i$, which contradicts the hypothesis that $\hat{s}_i$ improves on s_i. If $\hat{t} > t + 1$, we construct a strategy that agrees with $\hat{s}_i$ until $\hat{t} - 2$, and argue that it is as good a response as $\hat{s}_i$, and so on: The alleged sequence of improving deviations unravels from its endpoint. ■

What if the horizon is infinite? The proof above leaves open the possibility that player i could gain by some infinite sequence of deviations, even though he cannot gain by a single deviation in any subgame. Just as in dynamic programming, this possibility can be excluded if the payoff functions take the form of a discounted sum of per-period payoffs. More generally, the key condition is that the payoffs be "continuous at infinity." To make this precise, let h denote an infinite-horizon history, i.e., an

1. Even more precisely, there cannot be a history h^t such that the restriction of $\hat{s}_i$ to the subgame $G(h^t)$ is a better response than the restriction of s_i is.

outcome of the infinite-horizon game. For a fixed infinite-horizon history h, let h^t denote *the restriction of* h to the first t periods.

Definition 4.1 A game is *continuous at infinity* if for each player i the utility function u_i satisfies

$$\sup_{h, \tilde{h} \text{ s.t. } h^t = \tilde{h}^t} |u_i(h) - u_i(\tilde{h})| \to 0 \text{ as } t \to \infty.$$

This condition says that events in the distant future are relatively unimportant. It will be satisfied if the overall payoffs are a discounted sum of per-period payoffs $g_i^t(a^t)$ and the per-period payoffs are uniformly bounded, i.e, there is a B such that

$$\max_{t, a^t} |g_i^t(a^t)| < B.$$

Theorem 4.2 (one-stage deviation principle for infinite-horizon games) In an infinite-horizon multi-stage game with observed actions that is continuous at infinity, profile s is subgame perfect if and only if there is no player i and strategy $\hat{s}_i$ that agrees with s_i except at a single t and h^t, and such that $\hat{s}_i$ is a better response to s_{-i} than s_i conditional on history h^t being reached.

Proof The proof of the last theorem establishes necessity, and also shows that if s satisfies the one-stage-deviation condition then it cannot be improved by any finite sequence of deviations in any subgame. Suppose to the contrary that s were not subgame perfect. Then there would be a stage t and a history h^t where some player i could improve on his utility by using a different strategy $\hat{s}_i$ in the subgame starting at h^t. Let the amount of this improvement be $\varepsilon > 0$. Continuity at infinity implies that there is a t' such that the strategy s_i' that agrees with $\hat{s}_i$ at all stages before t' and agrees with s_i at all stages from t' on must improve on s_i by at least $\varepsilon/2$ in the subgame starting at h^t. But this contradicts the fact that no finite sequence of deviations can make any improvement at all. ■

This theorem and its proof are essentially the principle of optimality for discounted dynamic programming.

4.3 A First Look at Repeated Games†

4.3.1 The Repeated Prisoner's Dilemma

This section discusses the way in which repeated play introduces new equilibria by allowing players to condition their actions on the way their opponents played in previous periods. We begin with what is probably the best-known example of a repeated game: the celebrated "prisoner's dilemma," whose static version we discussed in chapter 1. Suppose that the per-period payoffs depend only on current actions ($g_i(a^t)$) and are as shown

	Cooperate	Defect
Cooperate	1,1	−1,2
Defect	2,−1	0,0

Figure 4.1

in figure 4.1, and suppose that the players discount future payoffs with a common discount factor δ. We will wish to consider how the equilibrium payoffs vary with the horizon T. To make the payoffs for different horizons comparable, we normalize to express them all in the units used for the per-period payoffs, so that the utility of a sequence $\{a^0, \ldots, a^T\}$ is

$$\frac{1-\delta}{1-\delta^{T+1}} \sum_{t=0}^{T} \delta^t g_i(a^t).$$

This is called the "average discounted payoff." Since the normalization is simply a rescaling, the normalized and present-value formulations represent the same preferences. The normalized versions make it easier to see what happens as the discount factor and the time horizon vary, by measuring all payoffs in terms of per-period averages. For example, the present value of a flow of 1 per period from date 0 to date T is $(1-\delta^{T+1})/(1-\delta)$; the average discounted value of this flow is simply 1.

We begin with the case in which the game is played only once. Then cooperating is strongly dominated, and the unique equilibrium is for both players to defect. If the game is repeated a finite number of times, subgame perfection requires both players to defect in the last period, and backward induction implies that the unique subgame-perfect equilibrium is for both players to defect in every period.[2]

If the game is played infinitely often, then "both defect every period" remains a subgame-perfect equilibrium. Moreover, it is the only equilibrium with the property that the play at each stage does not vary with the actions played at previous stages. However, if the horizon is infinite and $\delta > \frac{1}{2}$, then the following strategy profile is a subgame-perfect equilibrium as well: "Cooperate in the first period and continue to cooperate so long as no player has ever defected. If any player has ever defected, then defect for the rest of the game." With these strategies, there are two classes of subgames: class A, in which no player has defected, and class B, in which defect i on has occurred. If a player conforms to the strategies in every subgame in class A, his average discounted payoff is 1; if he deviates at time t and conforms to the (class B) strategies thereafter, his (normalized) payoff

2. This conclusion can be strengthened to hold for Nash equilibria as well. See section 5.2.

is

$$(1-\delta)(1+\delta+\cdots+\delta^{t-1}+2\delta^t+0+\cdots)=1-\delta^t(2\delta-1),$$

which is less than 1 as $\delta > \frac{1}{2}$. For any h^t in class B, the payoff from conforming to the strategies from period t on is 0; deviating once and then conforming gives -1 at period t and 0 in the future. Thus, in every subgame, no player can gain by deviating a single time from the specified strategy and then conforming, and so from the one-stage-deviation principle these strategies form a subgame-perfect equilibrium.

Depending on the size of the discount factor, there can be many other perfect equilibria. The next chapter presents the "folk theorem": Any feasible payoffs above the minmax levels (defined in chapter 5; in this example the minmax levels are 0) can be supported for a discount factor close enough to 1.[3] Thus, repeated play with patient players not only makes "cooperation"—meaning efficient payoffs—possible, it also leads to a large set of other equilibrium outcomes. Several methods have been proposed to reduce this multiplicity of equilibria; however, none of them has yet been widely accepted, and the problem remains a topic of research. We discuss one of the methods—"renegotiation-proofness"—in chapter 5.

Besides emphasizing the way that repeated play expands the set of equilibrium outcomes, the repeated prisoner's dilemma shows that the sets of equilibria of finite-horizon and infinite-horizon versions of the "same game" can be quite different, and in particular that new equilibria can arise when the horizon is allowed to be infinite. We return to this point at the end of this chapter.

4.3.2 A Finitely Repeated Game with Several Static Equilibria

The finitely repeated prisoner's dilemma has the same set of equilibria as the static version, but this is not always the case. Consider the multi-stage game corresponding to two repetitions of the stage game in figure 4.2. In the first stage of this game, players 1 and 2 simultaneously choose among

	L	M	R
U	0,0	3,4	6,0
M	4,3	0,0	0,0
D	0,6	0,0	5,5

Figure 4.2

3. The reason this holds only for large discount factors is that for small discount factors the short-term gain from deviating (for instance, deviating from cooperation in the prisoner's dilemma) necessarily exceeds any long-term losses that this behavior might create. See chapter 5.

U, M, D and L, M, R, respectively. At the end of the first stage the players observe the actions that were chosen, and in the second stage the players play the stage game again. As above, suppose that each player's payoff function in the multi-stage game is the discounted average of his or her payoffs in the two periods.

If this game is played once, there are three equilibria: (M, L), (U, M), and a mixed equilibrium ((3/7 U, 4/7 M), (3/7 L, 4/7 M)), with payoffs (4,3), (3,4), and (12/7, 12/7) respectively; the efficient payoff (5,5) is not attainable by an equilibrium. However, in the two-stage game, the following strategy profile is a subgame-perfect equilibrium if $\delta > 7/9$: "Play (D, R) in the first stage. If the first-stage outcome is (D, R), then play (M, L) in the second stage; if the first-stage outcome is not (D, R), then play ((3/7 U, 4/7 M), (3/7 L, 4/7 M)) in the second stage."

By construction, these strategies specify a Nash equilibrium in the second stage. Deviating in the first stage increases the current payoff by 1, and lowers the continuation payoffs for players 1 and 2 respectively from 4 or 3 to 12/7. Thus, player 1 will not deviate if $1 < (4 - 12/7)\delta$ or $\delta > 7/16$, and player 2 will not deviate if $1 < (3 - 12/7)\delta$ or $\delta > 7/9$.

4.4 The Rubinstein-Ståhl Bargaining Model[†]

In the model of Rubinstein 1982, two players must agree on how to share a pie of size 1. In periods 0, 2, 4, etc., player 1 proposes a sharing rule $(x, 1 - x)$ that player 2 can accept or reject. If player 2 accepts any offer, the game ends. If player 2 rejects player 1's offer in period $2k$, then in period $2k + 1$ player 2 can propose a sharing rule $(x, 1 - x)$ that player 1 can accept or reject. If player 1 accepts one of player 2's offers, the game ends; if he rejects, then he can make an offer in the subsequent period, and so on. This is an infinite-horizon game of perfect information. Note that the "stages" in our definition of a multi-stage game are not the same as "periods"—period 1 has two stages, corresponding to player 1's offer and player 2's acceptance or refusal.

We will specify that the payoffs if $(x, 1 - x)$ is accepted at date t are $(\delta_1^t x, \delta_2^t(1 - x))$, where x is players 1's share of the pie, and δ_1 and δ_2 are the two players' discount factors. (Rubinstein considered a somewhat larger class of preferences that allowed for a fixed per-period cost of bargaining in addition to the delay cost represented by the discount factors, and also allowed for utility functions that are not linear in the player's share of the pie.)

4.4.1 A Subgame-Perfect Equilibrium

Note that there are a great many *Nash* equilibria in this game. In particular, the strategy profile "player 1 always demands $x = 1$, and refuses all smaller shares; player 2 always offers $x = 1$ and accepts any offer" is a Nash equilibrium. However, this profile is not subgame perfect: If player 2 rejects player 1's first offer, and offers player 1 a share $x > \delta_1$, then player 1 should accept, because the best possible outcome if he rejects is to receive the entire pie tomorrow, which is worth only δ_1.

Here is a subgame-perfect equilibrium of this model: "Player i always demands a share $(1 - \delta_j)/(1 - \delta_i\delta_j)$ when it is his turn to make an offer. He accepts any share equal to or greater than $\delta_i(1 - \delta_j)/(1 - \delta_i\delta_j)$ and refuses any smaller share." Note that player i's demand of

$$\frac{1 - \delta_j}{1 - \delta_i\delta_j} = 1 - \frac{\delta_j(1 - \delta_i)}{1 - \delta_i\delta_j}$$

is the highest share for player i that is accepted by player j. Player i cannot gain by making a lower offer, for it too will be accepted. Making a higher (and rejected) offer and waiting to accept player j's offer next period hurts player i, as

$$\delta_i\left(1 - \frac{1 - \delta_i}{1 - \delta_i\delta_j}\right) = \delta_i^2\frac{1 - \delta_j}{1 - \delta_i\delta_j} < \frac{1 - \delta_j}{1 - \delta_i\delta_j}.$$

Similarly, it is optimal for player i to accept any offer of at least $\delta_i(1 - \delta_j)/(1 - \delta_i\delta_j)$ and to reject lower shares, since if he rejects he receives the share $(1 - \delta_i)/(1 - \delta_i\delta_j)$ next period.

Rubinstein's paper extends the work of Ståhl (1972), who considered a finite-horizon version of the same game. With a finite horizon, the game is easily solved by backward induction: The unique subgame-perfect equilibrium in the last period is for the player who makes the offer (let's assume it is player 1) to demand the whole pie, and for his opponent to accept this demand. In the period before, the last offerer (player 1) will refuse all offers that give him less than δ_1, for he can ensure $\delta_1 \cdot 1$ by refusing. And so on.

The finite-horizon model has two potential drawbacks relative to the infinite-horizon model. First, the solution depends on the length of the game, and on which player gets to make the last offer; however, this dependence becomes small as the number of periods grows to infinity, as is shown in exercise 4.5. Second, and more important, the assumption of a last period means that if the last offer is rejected the players are not allowed to continue to try to reach an agreement. In situations in which there is no outside opportunity and no per-period cost of bargaining, it is natural to assume that the players keep on bargaining as long as they haven't reached an agreement. Thus the only way to dismiss the suspicion that it matters whether one prohibits further bargaining after the exogenous finite horizon is to prove uniqueness in the infinite-horizon game.

4.4.2 Uniqueness of the Infinite-Horizon Equilibrium

Let us now demonstrate that the infinite-horizon bargaining game has a unique equilibrium. The following proof, by Shaked and Sutton (1984), uses the stationarity of the game to obtain an upper bound and a lower bound on each player's equilibrium payoff and then shows that the upper and lower bounds are equal. Section 4.6 gives an alternative proof of uniqueness that, although slightly longer, clarifies the uniqueness result through a generalization of the concept of iterative strict dominance.

To exploit the stationarity of the game, we define *the continuation payoffs of a strategy profile in a subgame starting at t* to be the utility in time-t units of the outcome induced by that profile. For example, the continuation payoff of player 1 at period 2 of a profile that leads to player 1's getting the whole pie at date 3 is δ_1, whereas this outcome has utility δ_1^3 in time-0 units.

Now we define $\underline{v}_1$ and $\bar{v}_1$ to be player 1's lowest and highest continuation payoffs of player 1 in any perfect equilibrium of any subgame that begins with player 1 making an offer. (More formally, $\underline{v}_1$ is the infimum or greatest lower bound of these payoffs, and $\bar{v}_1$ is the supremum.) Similarly, let $\underline{w}_1$ and $\bar{w}_1$ be player 1's lowest and highest perfect-equilibrium continuation payoffs in subgames that begin with an offer by player 2. Also, let $\underline{v}_2$ and $\bar{v}_2$ be player 2's lowest and highest perfect-equilibrium continuation payoffs in subgames beginning with an offer by player 2, and let $\underline{w}_2$ and $\bar{w}_2$ be player 2's lowest and highest perfect-equilibrium continuation payoffs in subgames beginning with an offer by player 1.

When player 1 makes an offer, player 2 will accept any x such that player 2's share (of $1 - x$) exceeds $\delta_2 \bar{v}_2$, since player 2 cannot expect more than $\bar{v}_2$ in the continuation game following his refusal. Hence, $\underline{v}_1 \geq 1 - \delta_2 \bar{v}_2$. By the symmetric argument, player 1 accepts all shares above $\delta_1 \bar{v}_1$, and $\underline{v}_2 \geq 1 - \delta_1 \bar{v}_1$.

Since player 2 will never offer player 1 a share greater than $\delta_1 \bar{v}_1$, player 1's continuation payoff when player 2 makes an offer, $\bar{w}_1$, is at most $\delta_1 \bar{v}_1$.

Since player 2 can obtain at least $\underline{v}_2$ in the continuation game by rejecting player 1's offer, player 2 will reject any x such that $1 - x < \delta_2 \underline{v}_2$. Therefore, player 1's highest equilibrium payoff when making an offer, $\bar{v}_1$, satisfies

$$\bar{v}_1 \leq \max(1 - \delta_2 \underline{v}_2, \delta_1 \bar{w}_1) \leq \max(1 - \delta_2 \underline{v}_2, \delta_1^2 \bar{v}_1).$$

Next, we claim that

$$\max(1 - \delta_2 \underline{v}_2, \delta_1^2 \bar{v}_1) = 1 - \delta_2 \underline{v}_2:$$

If not, then we would have $\bar{v}_1 \leq \delta_1^2 \bar{v}_1$, implying $\bar{v}_1 \leq 0$, but then $1 - \delta_2 \underline{v}_2 > \delta_1^2 \bar{v}_1$, as neither δ_2 nor $\underline{v}_2$ can exceed 1. Thus, $\bar{v}_1 \leq 1 - \delta_2 \underline{v}_2$. By symmetry, $\bar{v}_2 \leq 1 - \delta_1 \underline{v}_1$. Combining these inequalities, we have

$$\underline{v}_1 \geq 1 - \delta_2 \bar{v}_2 \geq 1 - \delta_2(1 - \delta_1 \underline{v}_1),$$

or

$$\underline{v}_1 \geq \frac{1 - \delta_2}{1 - \delta_1 \delta_2},$$

and

$$\bar{v}_1 \leq 1 - \delta_2(1 - \delta_1 \bar{v}_1),$$

or

$$\bar{v}_1 \leq \frac{1 - \delta_2}{1 - \delta_1 \delta_2};$$

because $\underline{v}_1 \leq \bar{v}_1$, this implies $\underline{v}_1 = \bar{v}_1$. Similarly,

$$\underline{v}_2 = \bar{v}_2 = \frac{1 - \delta_1}{1 - \delta_1 \delta_2},$$

$$\underline{w}_1 = \bar{w}_1 = \frac{\delta_1(1 - \delta_2)}{1 - \delta_1 \delta_2},$$

and

$$\underline{w}_2 = \bar{w}_2 = \frac{\delta_2(1 - \delta_1)}{1 - \delta_1 \delta_2}.$$

This shows that the perfect-equilibrium continuation payoffs are unique. To see that there is a unique perfect-equilibrium strategy profile, consider a subgame that begins with an offer by player 1. The argument above shows that player 1 must offer exactly $x = \underline{v}_1$. Although player 2 is indifferent between accepting and rejecting this offer, perfect equilibrium requires that he accept with probability 1: If player 2's strategy is to accept all $x < \underline{v}_1$ with probability 1, but to accept v_1 with probability less than 1, then no best response for player 1 exists. Hence, this randomization by player 2 is inconsistent with equilibrium. A similar argument applies in subgames that begin with an offer by player 2.

4.4.3 Comparative Statics

Note that as $\delta_1 \to 1$ for fixed δ_2, $v_1 \to 1$ and player 1 gets the whole pie, whereas player 2 gets the whole pie if $\delta_2 \to 1$ for fixed δ_1. Player 1 also gets the whole pie if $\delta_2 = 0$, since a myopic player 2 will accept any positive amount today rather than wait one period. Note also that even if $\delta_1 = 0$ player 2 does not get the whole pie if $\delta_2 < 1$: Due to his first-mover advantage, even a myopic player 1 receives a positive share. The first-mover advantage also explains why player 1 does better than player 2 even if the discount factors are equal: If $\delta_1 = \delta_2 = \delta$, then

$$v_1 = \frac{1}{1+\delta} > \tfrac{1}{2}.$$

As one would expect, this first-mover advantage disappears if we take the time periods to be arbitrarily short. To see this, let Δ denote the length of the time period, and set $\delta_1 = \exp(-r_1\Delta)$ and $\delta_2 = \exp(-r_2\Delta)$. Then, for Δ close to 0, δ_i is approximately $1 - r_i\Delta$, and v_1 converges to $r_2/(r_1 + r_2)$, so the relative patience of the players determines their shares. In particular, if $r_1 = r_2$ the players have equal shares in the limit. (See Binmore 1981 for a very thorough discussion of the Rubinstein-Ståhl model.)

4.5 Simple Timing Games[††]

4.5.1 Definition of Simple Timing Games

In a simple timing game, each player's only choice is when to choose the action "stop," and once a player stops he has no effect on future play. That is, if player i has not stopped at any $\tau < t$, his action set at t is

$$A_i(t) = \{\text{stop, don't stop}\};$$

if player i has stopped at some $\tau < t$, then $A_i(t)$ is the null action "don't move." Few situations can be exactly described in this way, because players typically have a wider range of choices. (For example, firms typically do not simply choose a time to enter a market; they also decide on the scale of entry, the quality level, etc.) But economists often abstract away from such details to study the timing question in isolation.

We will consider only two-player timing games, and restrict our attention to the subgame-perfect equilibria. Once one player has stopped, the remaining player faces a maximization problem that is easily solved. Thus, when considering subgame-perfect equilibria, we can first "fold back" subgames where one player has stopped and then proceed to subgames where neither player has yet stopped.[4] This allows us to express both players' payoffs as functions of the time

$$\hat{t} = \min\{t \mid a_i^t = \text{stop for at least one } i\}$$

at which the first player stops (with the strategies we will consider, this minimum is well defined); if no player ever stops, we set $\hat{t} = \infty$. We describe these payoffs using the functions L_i, F_i, and B_i: If only player i stops at $\hat{t}$, then player i is the "leader"; he receives $L_i(\hat{t})$, and his opponent receives "follower" payoff $F_j(\hat{t})$. If both players stop simultaneously at $\hat{t}$, the payoffs are $B_1(\hat{t})$ and $B_2(\hat{t})$. We will assume that

4. Although the one-player maximization problem will typically have a unique solution, this need not be the case. If there are multiple solutions, then one must consider the implications of each of them.

$$\lim_{\hat{t} \to +\infty} L_i(\hat{t}) = \lim_{\hat{t} \to +\infty} F_i(\hat{t}) = \lim_{\hat{t} \to +\infty} B_i(\hat{t}),$$

which will be the case if payoffs are discounted.

The last step in describing these games is to specify the strategy spaces. We begin with the technically simpler case where time is discrete, as it has been in our development so far. Since the feasible actions at each date until some player stops are simply {stop, don't stop}, and since once a player stops the game effectively ends (remember that we have folded back any subsequent play), the history at date t is simply the fact that the game is still going on then. Thus, pure strategies s_i are simply maps from the set of dates t to {stop, don't stop}, behavior strategies b_i specify a conditional probability $b_i(t)$ of stopping at t if no player has stopped before, and mixed strategies σ_i are probability distributions over the pure strategies s_i.

For some games, the set of equilibria is easier to compute in a model where time is continuous. The pure strategies, as in discrete time, are simply maps from times t to {stop, don't stop}, but two complications arise in dealing with mixed strategies. First, the formal notion of behavior strategies becomes problematic when players have a continuum of information sets. (This was first noted by Aumann (1964).) We will sidestep the question of behavior strategies by working only with the mixed (i.e., strategic-form) strategies. The second problem is that, as we will see, the set of continuous-time mixed strategies as they are usually defined is too small to ensure that the continuous-time model will capture the limits of discrete-time equilibria with short time intervals, although it does capture the short-time-interval limits of some classes of games.

Putting these problems aside, we introduce the space of continuous-time mixed strategies that we will use in most of this section. Given that pure strategies are stopping times, it is natural to identify mixed strategies as cumulative distribution functions G_i on $[0, \infty)$. In other words, $G_i(t)$ is the probability that player i stops at or before time t. (To be cumulative distribution functions, the G_i must take values in the interval $[0, 1]$ and be nondecreasing and right-continuous.[5]) The functions G_i need not be continuous; let

$$\alpha_i(t) = G_i(t) - \lim_{\tau \uparrow t} G_i(\tau)$$

be the size of the jump at t. When $\alpha_i(t)$ is nonzero, player i stops with probability $\alpha_i(t)$ at exactly time t; this is called an "atom" of the probability distribution. Where G_i is differentiable, its derivative dG_i is the probability density function; the probability that player i stops between times t and

5. The function $G(\cdot)$ is right-continuous at t if

$$\lim_{\tau \downarrow t} G(\tau) = G(t).$$

It is right-continuous if it is right-continuous at each t.

$t + \varepsilon$ is approximately $\varepsilon dG_i(t)$. Player i's payoff function is then

$$u_i(G_1, G_2) = \int_0^t [L_i(s)(1 - G_j(s))\,dG_i(s) + F_i(s)(1 - G_i(s))\,dG_j(s)] + \sum_s \alpha_i(s)\alpha_j(s)B_i(s).$$

That is, there is probability $dG_i(s)$ that player i stops at date s. If player j hasn't stopped yet, which has probability $1 - G_j(s)$, player i is the leader and obtains $L_i(s)$. And similarly for the other terms.

We now develop two familiar games of timing: the war of attrition and the preemption game.[6]

4.5.2 The War of Attrition

A classic example of a timing game is the war of attrition, first analyzed by Maynard Smith (1974).[7]

Stationary War of Attrition

In the *discrete-time* version of the stationary war of attrition, two animals are fighting for a prize whose current value at any time $t = 0, 1, \ldots$ is $v > 1$; fighting costs 1 unit per period. If one animal stops fighting in period t, his opponent wins the prize without incurring a fighting cost that period, and the choice of the second stopping time is irrelevant. If we introduce a per-period discount factor δ, the (symmetric) payoff functions are

$$L(\hat{t}) = -(1 + \delta + \cdots + \delta^{\hat{t}-1}) = -\frac{1 - \delta^{\hat{t}}}{1 - \delta}$$

and

$$F(\hat{t}) = -(1 + \delta + \cdots + \delta^{\hat{t}-1}) + \delta^{\hat{t}}v = L(\hat{t}) + \delta^{\hat{t}}v.$$

If both animals stop simultaneously, we specify that neither wins the prize, so that

$$B_1(\hat{t}) = B_2(\hat{t}) = L(\hat{t}).$$

(Exercise 4.1 asks you to check that other specifications with $B(\hat{t}) < F(\hat{t})$ yield similar conclusions when the time periods are sufficiently short.) Figure 4.3 depicts $L(\cdot)$ and $F(\cdot)$ for the continuous-time version of this game.

This stationary game has several Nash equilibria. Here is one: Player 1's strategy is "never stop" and player 2's is "always stop." There is a unique

6. See Katz and Shapiro 1986 for hybrids of these two games.

7. Another name for the war of attrition is "chicken." The classic game of chicken is played in automobiles. In one version the cars head toward a cliff side by side, and the first driver to stop loses; in the other version the cars head toward each other, and the first driver to swerve out of the way to avoid the collision loses. We do not recommend experimental studies of either version of this game.

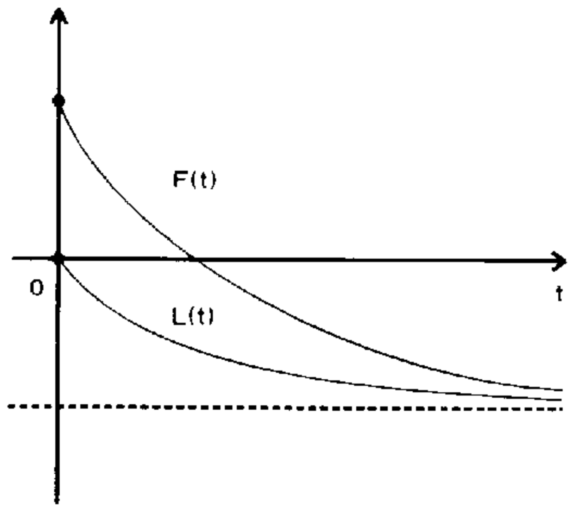

Figure 4.3

symmetric equilibrium, which is stationary and involves mixed strategies: For any p, let "always p" be the behavior strategy "if the other player has not stopped before t, then stop at t with probability p." The equivalent strategic-form mixed strategy assigns probability $(1 - p)^t p$ to the pure strategy "stop at t if the other player hasn't stopped before then." For the stationary symmetric profile (always p, always p) to be an equilibrium, it is necessary that, for any t, the payoff to stopping at t conditional on the opponent's not having stopped previously, which is $L(t)$, is equal to

$$p[F(t)] + (1 - p)[L(t + 1)],$$

the payoff to staying in until $t + 1$ and dropping out then unless the opponent drops out today. (If the opponent does drop out before t, then the strategies "stop at t" and "stop at $t + 1$" yield the same payoff.) Equating these terms gives $p^* = 1/(1 + v)$, which ranges from 1 to 0 as v ranges from 0 to infinity. Another way to arrive at this conclusion is to note that by staying in for one more period, a player gains v with probability p and loses the fighting cost 1 during that period with probability $1 - p$. For him to be indifferent between staying in for one more period and stopping now, it must be the case that $pv = 1 - p$, which yields the above expression for p^*.

Thus, "always p^*" is the only candidate for a stationary symmetric equilibrium. To check that it is indeed an equilibrium, note that if player 1 plays "always p^*" the payoff to each possible stopping time t for player 2 is 0.

At this stage, the reader may wonder whether the Nash equilibria are subgame perfect. If the players are free to quit when they want and are not committed to abide by their date-0 choice of stopping time, do they want to deviate? The answer is No: All stationary Nash equilibria (i.e. equilibria with strategies that are independent of calendar time) are subgame perfect.

(To see this, note that the stationarity of the payoffs implies that all subgames where both players are still active are isomorphic.)

The *continuous-time* formulation is very convenient for wars of attrition. Consider the continuous-time version of the example considered above, where the terms δ^t are replaced by $\exp(-rt)$ and r is the rate of interest. Let $G_i(t)$ denote the probability that player i stops at or before t (that is, $G_i(\cdot)$ is a cumulative distribution function). As in the discrete-time version of the game, there is a stationary symmetric equilibrium G with the property that at each date the players are indifferent between stopping at time t and waiting a bit longer, until $t + \varepsilon$, to see if the opponent stops first. Conditional on not stopping before t, to the first order in ε,[8] the marginal cost of waiting ε longer is ε and the expected reward from doing so is $v\,dG/(1 - G)$. Equating these terms yields $dG/(1 - G) = 1/v$, so that G is the exponential distribution $G^*(t) = 1 - \exp(-t/v)$. (As in the discrete-time case, to verify that these are equilibrium strategies, note that if palyer 1 uses G^* then player 2's expected payoff to any strategy is 0.)

Thus, the war of attrition does have a symmetric equilibrium in the kind of continuous-time strategies we introduced above. Moreover, this equilibrium is the limit of the symmetric equilibria of the discrete-time game as the interval Δ between periods goes to 0, as we will now show. To make the discrete- and continuous-time formulations comparable, we assume that fighting costs 1 per unit of *real* time. Hence, if in discrete time the real length of each period is Δ (so that there are $1/\Delta$ periods per unit of time), the fighting cost is Δ per period. The value of the prize v does not need to be adjusted when the period length changes, as v was taken to be a stock rather than a flow in both formulations. The discrete-time equilibrium strategy is now given by $p^*v = (1 - p^*)\Delta$ or $p^* = \Delta/(\Delta + v)$. Fix a real time t, and let $n = t/\Delta$ be the number of (discrete-time) periods between 0 and t. The probability that a player does not stop before t in the discrete-time formulation is

$$1 - G(t) = (1 - p^*)^n = \left(\frac{v}{v + \Delta}\right)^{t/\Delta} = \exp\left[-\frac{t}{\Delta}\ln\left(1 + \frac{\Delta}{v}\right)\right]$$
$$\simeq \exp\left(-\frac{t}{\Delta}\frac{\Delta}{v}\right) = \exp\left(-\frac{t}{v}\right) \text{ for } \Delta \text{ small.}$$

Thus, the symmetric discrete-time equilibrium converges to the symmetric continuous-time equilibrium when Δ tends to 0.

Nonstationary Wars of Attrition

More generally, games satisfying the following (discrete- or continuous-time) conditions can be viewed as wars of attrition: For all players i and

8. We will say that a term $f(\varepsilon)$ is not of order ε if $\lim_{\varepsilon \to 0} f(\varepsilon)/\varepsilon = 0$.

all dates t,

(i) $F_i(t) \geq F_i(\tau)$ for $\tau > t$,

(ii) $F_i(t) \geq L_i(\tau)$ for $\tau > t$,

(iii) $L_i(t) = B_i(t)$,

(iv) $L_i(0) > L_i(+\infty)$,

(v) $L_i(+\infty) = F_i(+\infty)$.

Condition i states that if player i's opponent is going to stop first in the subgame starting at t, then player i prefers that his opponent stop immediately at t. Condition ii says that each player i prefers his opponent stopping first at any time t to any outcome where player i stops first at some $\tau > t$. The motivation for this condition is that if player j stops at t, player i can always stay in until $\tau > t$ and quit at τ and obtain $L_i(\tau)$ plus the economized fighting costs between t and τ.

Condition iii asserts that when a player stops, it does not matter if the opponent stops or stays; this assumption simplifies the study of the discrete-time formulation and is irrelevant under continuous time. Condition iv states that fighting forever is costly—each player would rather quit immediately than fight forever. Condition v is automatically satisfied if players discount their payoffs and the payoffs are bounded.

Two nonstationary variants satisfying these conditions and stronger assumptions have appeared frequently in the literature: "eventual continuation" and "eventual stopping." (We state these further assumptions for the discrete-time framework in order not to discuss continuous-time strategies.) In either case, subgame perfection uniquely pins down equilibrium behavior.

Eventual Continuation Make assumptions i–v and the following additional assumptions:

(ii′) $F_i(t+1) > L_i(t)$ for all i and t.

(vi) For all i, there exists $T_i > 0$ such that $L_i(t) > L_i(+\infty)$ for $t < T_i$ and $L_i(t) < L_i(+\infty)$ for $t > T_i$.

(vii) For all i, there exists $\tilde{T}_i$ such that $L_i(\cdot)$ is strictly decreasing before $\tilde{T}_i$ and increasing after $\tilde{T}_i$.

Condition ii′, which states that fighting for one period is worthwhile if successful, is a strengthening of condition ii. (To see that conditions i and ii′ imply condition ii, note that for $\tau > t$, $F_i(t) \geq F_i(\tau + 1) > L_i(\tau)$.) Condition vi states that even though at the start of the game it is better to quit than to fight forever, things get better later on so that, ignoring past sunk costs, the player would rather continue fighting than quit. Condition vii states that $L_i(\cdot)$ has a single peak. Note that, necessarily, $\tilde{T}_i \geq T_i$. In an industrial-organization context, conditions vi and vii correspond to the market growing or the technology improving over time (for exogenous reasons or because of learning by doing).

Example of eventual continuation Fudenberg et al. (1983) study an example of eventual continuation: Two firms are engaged in a patent race, and "stopping" means abandoning the race. The expected productivity of research is initially low, so that if both firms do R&D until one of them makes a discovery then both firms have a negative expected value. However, the productivity of R&D increases over time, so that there are times T_1 and T_2 such that, if both firms are still active at T_i, then it is a dominant strategy for firm i to never stop.

For simplicity, we give the continuous-time version of the patent-race game. Suppose that the patent has value v. If firm i has not quit before date t, it pays $c_i\,dt$ and makes a discovery with probability $x_i(t)\,dt$ between t and $t + dt$. The instantaneous flow profit is thus $[x_i(t)v - c_i]\,dt$. (The probability that firm j discovers between t and $t + dt$ is infinitesimal.) Suppose that $dx_i/dt > 0$ (due to learning).

In this game,

$$L_i(t) = \int_0^t [x_i(\tau)v - c_i] \exp\left(- \int_0^\tau [x_1(s) + x_2(s)]\,ds \right) \exp(-r\tau)\,d\tau,$$

where r is the rate of interest. The probability that no one has discovered at date τ conditional on both players having stayed in the race is

$$\exp\left(- \int_0^\tau [x_1(s) + x_2(s)]\,ds \right).$$

We assume that an R&D monopoly is viable:

$$0 < \int_0^\infty [x_i(\tau)v - c_i] \exp\left(- \int_0^\tau x_i(s)\,ds \right) \exp(-r\tau)\,d\tau = F_i(0),$$

and that a duopoly is not viable at date 0: $L_i(\infty) < 0$. (Recall that $L_i(\infty)$ is the date-0 payoff if neither firm ever stops.) Because $x_i(\cdot)$ is increasing, if a monopoly is viable at date 0 then it is viable from any date $t > 0$ on.[9] Therefore, it is optimal for each player to stay in until discovery once his opponent has quit. The follower's payoff is thus

$$\begin{aligned} F_i(t) = {} & \int_0^t [x_i(\tau)v - c_i] \exp\left(- \int_0^\tau [x_1(s) + x_2(s)]\,ds \right) \exp(-r\tau)\,d\tau \\ & + \int_t^\infty [x_i(\tau)v - c_i] \exp\left[- \left(\int_0^\tau x_i(s)\,ds + \int_0^t x_j(s)\,ds \right) \right] \\ & \times \exp(-r\tau)\,d\tau. \end{aligned}$$

9. Note also that $F_i(0) > 0$ and $x_i(\cdot)$ increasing imply that there exists a time such that $x_i(t)v > c_i$ after that time. Therefore L_i is first decreasing and then increasing. The time $\tilde{T}_i$ defined in condition vii is given by the equation $x_i(\tilde{T}_i)v = c_i$.

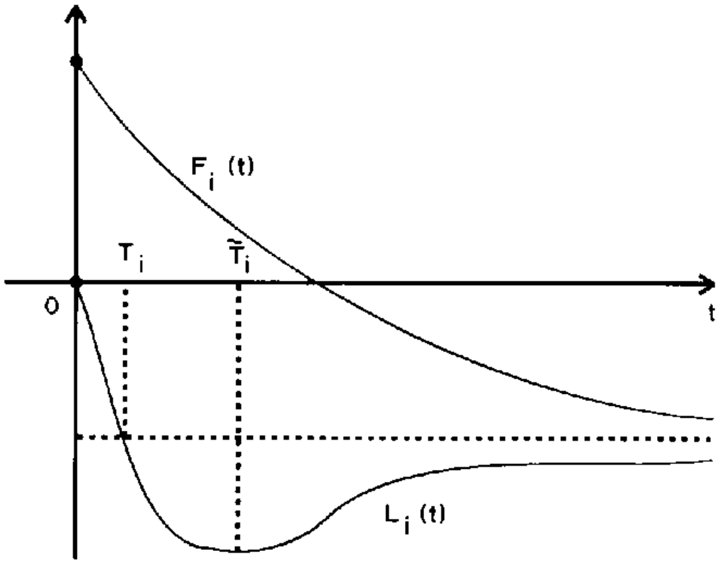

Figure 4.4

The leader's and the follower's payoffs in the continuous-time patent race are depicted in figure 4.4. It is clear that assumptions i–v, ii′, vi, and vii are satisfied in the discrete-time version of the game as long as the interval between the periods is small (i.e., when discrete time is close to continuous time).

Uniqueness under eventual continuation Condition vi guarantees that player i never quits after T_i: By quitting at date $t > T_i$ he obtains

$$L_i(t) < L_i(+\infty) = F_i(+\infty) \leq F_i(\tau)$$

for all τ. Thus, he always gets more by never quitting. Thus, quitting is a (conditionally) strictly dominated strategy at date $t > T_i$. Let us now assume that the times (T_1, T_2) defined in condition vi satisfy $T_1 + 1 < T_2$. If the time interval between the periods is short and the players are not quite the same (for instance, if $c_1 \neq c_2$ or $x_1(\cdot) \neq x_2(\cdot)$ in the patent race), this condition (or the symmetric condition) is likely to be satisfied. We claim that the war of attrition has a unique equilibrium, and that in this equilibrium player 2 quits at date 0 (there is no war).

Uniqueness is proved by backward induction. At date $T_1 + 1$, if both players are still fighting, player 2 knows that player 1 will never quit. Because $T_1 + 1 < T_2$,

$$L_2(T_1 + 1) > L_2(+\infty).$$

Furthermore, because $L_2(\cdot)$ has a single peak (condition vii),

$$L_2(T_1 + 1) > L_2(t)$$

for all $t > T_1 + 1$. Hence, it is optimal for player 2 to quit at $T_1 + 1$. Consider now date T_1. Because $F_1(T_1 + 1) > L_1(T_1)$ (condition ii′), player

1 does not quit. And by the same reasoning as before, $L_2(T_1) > L_2(t)$ for all $t > T_1$. Hence, player 2 quits at T_1 if both players are still around at that date. The same reasoning shows that player 2 quits and player 1 stays in at any date $t < T_1$. There exists a unique subgame-perfect equilibrium.

One reason why $T_1 + 1$ might be less than T_2 in the patent-race example is that firm 1 entered the patent race $k \geq 2$ periods before firm 2. Then, if the two firms have the same technology ($x_2(t) = x_1(t - k)$ and $c_1 = c_2$),

$$T_1 = T_2 - k.$$

If periods are short, then the case $T_1 = T_2 - 2$ seems like a small advantage for player 1, yet it is sufficient to make firm 1 the "winner" without a fight. In the terminology of Dasgupta and Stiglitz (1980), this game exhibits "ε-preemption," as an ε advantage proves decisive. Harris and Vickers (1985) develop related ε-preemption arguments. Hendricks and Wilson (1989) characterize the equilibria of a large family of discrete-time wars of attrition; Hendricks et al. (1988) do the same for the continuous-time version.

Eventual Stopping Make assumptions i–v and the following additional assumptions:

(viii) There exists $T_2 > 0$ such that $\forall t < T_2, L_2(t) < F_2(t), \forall t > T_2, F_2(t) \leq L_2(t)$, and $\forall t \leq T_2$, $F_1(t + 1) > L_1(t)$.

(ix) For all i, $L_i(\cdot)$ has a single peak. It increases strictly before some time $\tilde{T}_i$ and decreases strictly after time $\tilde{T}_i$. Furthermore, $\tilde{T}_2 \leq \tilde{T}_1$.

Assumption viii states that after some date T_2, player 2 is better off exiting than staying even if the other player has quit. The following example illustrates these conditions. Note that necessarily $\tilde{T}_2 < T_2$.

Example of eventual stopping As with eventual continuation, we give the example and draw the payoff functions in continuous time; the proof of uniqueness is performed in the discrete-time framework. Two firms wage duopoly competition in a market. If one quits, the other becomes a monopoly. Suppose that the firms differ only in their flow fixed cost, $f_1 < f_2$. The gross flow profits are $\Pi^m(t)$ for a monopolist and $\Pi^d(t)$ for a duopolist, where $\Pi^m(t) > \Pi^d(t)$ for all t. Demand is declining, so $\Pi^m(\cdot)$ and $\Pi^d(\cdot)$ are strictly decreasing. Suppose that there exist $\tilde{T}_2$ and T_2 such that $0 < \tilde{T}_2 < T_2 < +\infty$, $\Pi^d(\tilde{T}_2) = f_2$ (firm 2 stops making profit as a duopolist at date $\tilde{T}_2$), and $\Pi^m(T_2) = f_2$ (firm 2 is no longer viable as a monopolist after date T_2). The payoffs $F_2(\cdot)$ and $L_2(\cdot)$ are represented in figure 4.5 (firm 1's payoffs have similar shapes). After T_2 it is optimal for firm 2 to quit immediately, so in continuous time $F_2(t) = L_2(t)$ for $t > T_2$. (In discrete time and with short time periods, $F_2(t)$ is slightly lower than $L_2(t)$, since firm 2, as a follower, stays one period longer than the leader and loses money during that period.) In discrete time, this example satisfies assump-

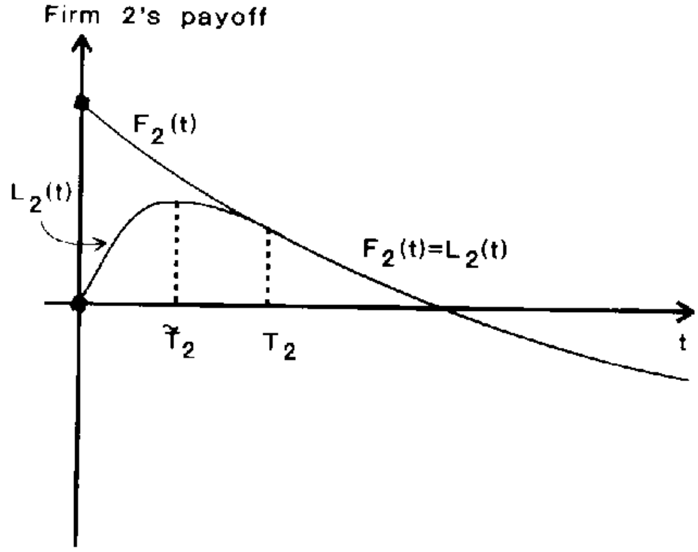

Figure 4.5

tions i–v, viii, and ix for sufficiently short time intervals between the periods.

Uniqueness under Eventual Stopping First we claim that player 2 quits at any date $t > T_2$. By stopping it gets $L_2(t)$. Because $L_2(t) \geq F_2(t)$ from condition viii, and because $L_2(\cdot)$ is strictly decreasing from condition ix, $L_2(t) > L_2(\tau) \geq F_2(\tau)$ for $\tau > t$. Hence, it is a strictly dominant strategy for player 2 to quit at any $t > T_2$. Therefore, $F_1(T_2 + 1) > L_1(T_2)$ implies that player 1 stays in at T_2. Because $L_2(T_2) > L_2(T_2 + 1)$, player 2 quits at T_2. The same holds by induction for any t greater than $\tilde{T}_2$. Before $\tilde{T}_2$, neither player quits since $L_i(\cdot)$ is increasing. Hence, the unique equilibrium of the game has player 2 quit first at $\tilde{T}_2$, and the two players' payoffs are $F_1(\tilde{T}_2)$ and $L_2(\tilde{T}_2)$.

Ghemawat and Nalebuff (1985) and Fudenberg and Tirole (1986) give examples of declining industries that fit the eventual-stopping example. Ghemawat and Nalebuff argue that if exit is an all-or-nothing choice, as in the simple model we have been considering, then a big firm will become unprofitable before a smaller one will, and so the big firm will be forced to exit first. Moreover, foreseeing this eventual exit, the small firm will stay in, and backward induction implies that the big firm exits once the market shrinks enough that staying in earns negative flow profits. (Whinston (1986) shows that this conclusion need not obtain if the big firm is allowed to shed capacity in small units.)

4.5.3 Preemption Games

Preemption games are a rough opposite to the war of attrition, with $L(\hat{t}) > F(\hat{t})$ for some range of times $\hat{t}$. Here the specification of the payoff to simultaneous stopping, $B(\cdot)$, is more important than in the war of attrition, as if L exceeds F we might expect both players to stop simultaneously. One example of a preemption game is the decision of whether and when to build a new plant or adopt a new innovation when the market is big enough to support only one such addition (Reinganum 1981a,b; Fudenberg and Tirole 1985). In this case $B(t)$ is often less than $F(t)$, as it can be better to let an opponent have a monopoly than to incur duopoly losses.

One very stylized preemption game is "grab the dollar." In this stationary game, time is discrete ($t = 0, 1, \ldots$) and there is a dollar on the table, which either or both of the players can try to grab. If only one player grabs, that player receives 1 and the other 0; if both try to grab at once, the dollar is destroyed and both pay a fine of 1; if neither player grabs, the dollar remains on the table. The players use the common discount factor δ, so that $L(t) = \delta^t$, $F(t) = 0$, and $B(t) = -\delta^t$ for all t. Like the war of attrition, this game has asymmetric equilibria, where one player "wins" with probability 1, and also a symmetric mixed-strategy equilibrium, where each player grabs the dollar with probability $p^* = \frac{1}{2}$ in each period. (It is easy to check that this yields a symmetric equilibrium; to see that it is the only one, note that each player must be indifferent between stopping—i.e., grabbing—at date t, which yields payoff $\delta^t((1 - p^*(t)) - p^*(t))$ if the other has not stopped before date t and 0 otherwise, and never stopping, which yields payoff 0, so that $p^*(t)$ must equal $\frac{1}{2}$ for all t.) The payoffs in the symmetric equilibrium are $(0, 0)$, and the distribution over outcomes is that the probability that player 1 alone stops first at t, the probability that player 2 alone stops first at t, and the probability that both players stop simultaneously at t are all equal to $(\frac{1}{4})^{t+1}$. Note that these probabilities are independent of the per-period discount factor, δ, and thus of the period length, Δ, in contrast to the war of attrition, where the probabilities were proportional to the period length. This makes finding a continuous-time representation of this game more difficult.

To understand the difficulties, let t denote a fixed "real time" after the initial time 0, define the number of periods between time 0 and time t when the real time length of the period is Δ as $n(t, \Delta) = t/\Delta$, and consider what happens as $\Delta \to 0$. The probability that at least one player has stopped by t is $1 - (\frac{1}{4})^{n(t,\Delta)}$, which converges to 1 as $\Delta \to 0$. The limit of the equilibrium distribution over outcomes is probability $\frac{1}{3}$ that player 1 wins the dollar at time 0, probability $\frac{1}{3}$ that player 2 wins at time 0, probability $\frac{1}{3}$ that both grab simultaneously at time 0, and probability 0 that the game continues beyond time 0. Fudenberg and Tirole (1985) observed that this limiting distribution cannot be expressed as an equilibrium in continuous-time strategies of the kind we have considered so far: If the game ends with probability 1 at time 0, then, for at least one player i, $G_i(0)$ must equal 1; but then there would be probability 0 that player i's opponent wins the dollar. The problem is that different sequences of discrete-time strategy profiles converge to a limit in which the game ends with probability 1 at the start, including "stop with probability 1 at time 0" and "stop with conditional probability $p > 0$ at each period." The usual continuous-time strategies implicitly associate an atom of size 1 in continuous time with an atom of that size in discrete time, and this cannot represent the limit of the discrete-time equilibria, where the atom at time 0 corresponds not to probability 1 of stopping at exactly time 0 but rather to an "interval of atoms" at all times just after 0. We proposed an extended version of the continuous-time strategies and payoff functions to capture the limit of discrete-time equilibria in the particular preemption game we were analyzing. Simon (1988) has developed a related, more general approach and applied it to a broader class of games.

As another example of a preemption game, suppose that $L(t) = 14 - (t - 7)^2$, $F(t) = 0$, and $B(t) < 0$. These payoffs are meant to describe a situation where either of two firms can introduce a new product. The product will have no effect on their existing business, which is why $F(t) = 0$, and the combination of fixed costs and aggressive duopoly pricing implies that both firms will lose money if both develop the product. Thus, once one firm introduces the product the other firm never will, and the only case in which both firms would introduce is if they did so simultaneously.

To avoid the need to consider the possibility of such mistakes and the associated mixed strategies, let us follow Gilbert and Harris (1984) and Harris and Vickers (1985) and make the simplifying assumption that player 1 can stop only in even-numbered periods ($t = 0, 2, \ldots$) and player 2 can stop only in odd-numbered periods ($t = 1, 3, \ldots$), so that the game is one of perfect information. There are three Pareto-efficient outcomes for the players—either player 1 stops at $t = 6$ or $t = 8$ or player 2 stops at $t = 7$—with two Pareto-efficient payoffs: $(13, 0)$ and $(0, 14)$. In the unique subgame-perfect equilibrium, firm 1 stops at $t = 4$, which is the first t where $L(t) > F(t)$. This is an example of "rent dissipation": Although there are possible rents to be made from introducing the product later, in equilibrium the race to be first forces the introduction of the product at the first time when rents are nonnegative.

4.6 Iterated Conditional Dominance and the Rubinstein Bargaining Game†††

The last two sections presented several examples of infinite-horizon games with unique equilibria. The uniqueness arguments there can be strengthened, in that these games have a unique profile that satisfies a weaker concept than subgame perfection.

Definition 4.2 In a multi-stage game with observed actions, action a_i^t is *conditionally dominated at stage t given history h^t* if, in the subgame beginning at h^t, every strategy for player i that assigns positive probability to a_i^t is strictly dominated. Iterated conditional dominance is the process that, at each round, deletes every conditionally dominated action in every subgame, given the opponents' strategies that have survived the previous rounds.

It is easy to check that iterated conditional dominance coincides with subgame perfection in finite games of perfect information. In these games it also coincides with Pearce's (1984) extensive-form rationalizability. In general multi-stage games, any action ruled out by iterated conditional dominance is also ruled out by extensive-form rationalizability, but the exact relationship between the two concepts has not been determined.

In a game of imperfect information, iterated conditional dominance can be weaker than subgame perfection, as it does not assume that players forecast that an equilibrium will occur in every subgame. To illustrate this point, consider a one-stage, simultaneous-move game. Then iterated conditional dominance coincides with iterated strict dominance, subgame perfection coincides with Nash equilibrium, and iterated strict dominance is in general weaker than Nash equilibrium.

Theorem 4.3 In a finite- or infinite-horizon game of perfect information, no subgame-perfect strategy profile is removed by iterated conditional dominance.

Proof Proving this theorem is exercise 4.7.

Iterated conditional dominance is quite weak in some games. For example, in a repeated game with discount factor near 1, no action is conditionally dominated. (It is always worthwhile to play any fixed action today if playing the action induces the opponents to play cooperatively in every future period and if not playing the action causes the opponents to play something you do not like.) However, in games that "end" when certain actions are played, iterated conditional dominance has more bite, since if a player's current action ends the game then his opponents will not be able to "punish" him in the future. One such example is the infinite-horizon version of the bargaining model in section 4.4.

Let us see how iterated conditional dominance gives a unique solution. Note first that a player never accepts an offer that gives him a negative share (i.e., such that he loses money): Accepting such an offer is strictly dominated by the strategy "reject any offer, including this one, and make only offers that, if accepted, give a positive share." Next, in any subgame where player 2 has just made an offer, it is conditionally dominated for player 1 to refuse if the offer has player 1's share of the pie, x, exceeding δ_1; similarly, player 2 must accept any $x < 1 - \delta_2$. These are all the actions that are removed at the first round of iteration. At the second round, we conclude that

(i) player 2 will never offer player 1 more than δ_1,

(ii) player 2 will reject any $x > 1 - \delta_2(1 - \delta_1)$, because he can get $\delta_2(1 - \delta_1)$ by waiting one period,

(iii) player 1 never offers $x < 1 - \delta_2$, and

(iv) player 1 rejects all $x < \delta_1(1 - \delta_2)$.

To continue on, imagine that, after k rounds of elimination of conditionally dominated strategies, we have concluded that player 1 accepts all $x > x^k$, and player 2 accepts all $x < y^k$, with $x^k > y^k$. Then, after one more round, we conclude that

(i) player 2 never offers $x > x^k$,

(ii) player 2 rejects all $x > 1 - \delta_2(1 - x^k)$,

(iii) player 1 never offers $x < y^k$, and

(iv) player 1 rejects all $x < \delta_1 y^k$.

At the next round of iteration, we claim that player 1 must accept all $x > x^{k+1} = \delta_1(1 - \delta_2) + \delta_1\delta_2 x^k$, and that player 2 accepts all $x < y^{k+1} = 1 - \delta_2 + \delta_1\delta_2 y^k$, where $x^{k+1} > y^{k+1}$. We will check this claim for player 1: If player 1 refuses an offer of player 2 in some subgame, one of three things can happen. Either (a) no agreement is ever reached, which has payoff 0, or (b) player 2 accepts one of player 1's offers, which has a current value of at most $\delta_1(1 - \delta_2(1 - x^k))$ (since the soonest this can happen is next period, and player 2 refuses x above $(1 - \delta_2(1 - x^k))$), or (c) player 1 accepts one of player 2's offers, which has a current value of at most $\delta_1^2 x^k$. Simple algebra shows that, for all discount factors δ_1, δ_2, the payoff in case b is largest, so player 1 accepts all $x > \delta_1(1 - \delta_2) + \delta_1\delta_2 x^k$.

The x^k and y^k are monotonic sequences, with limits $x^\infty = \delta_1(1 - \delta_2)/(1 - \delta_1\delta_2)$ and $y^\infty = (1 - \delta_2)/(1 - \delta_1\delta_2)$. Iterated conditional dominance shows that player 2 rejects all $x > 1 - \delta_2(1 - x^\infty) = y^\infty$ and accepts all $x < y^\infty$, so the unique equilibrium outcome is for player 1 to offer exactly y^∞ and for player 2 to accept. (There is no equilibrium profile in which player 2 rejects y^∞ with positive probability, for then player 1 would want to offer "just below" y^∞, so that no best response for player 1 exists.)

4.7 Open-Loop and Closed-Loop Equilibria††

4.7.1 Definitions

The terms *closed-loop* and *open-loop* are used to distinguish between two different information structures in multi-stage games. Our definition of a multi-stage game with observed actions corresponds to the closed-loop information structure, where players can condition their play at time t on the history of play until that date. In the terminology of the literature on optimal control, the corresponding strategies are called *closed-loop strate-*

gies or feedback strategies, while *open-loop* strategies are functions of calendar time alone.

Determining which are the appropriate strategies to consider is the same as determining the information structure of the game. Suppose first that the players never observe any history other than their own moves and time, or that at the beginning of the game they must choose time paths of actions that depend only on calendar time. (These two situations are equivalent from the extensive-form viewpoint, as the role of information sets is to describe what information players can use in choosing their actions.) In this case all strategies are open-loop, and all Nash equilibria (which in this case coincide with perfect equilibria) are in open-loop strategies. An equilibrium in open-loop strategies is called an *open-loop equilibrium*. (As with "Cournot" and "Stackelberg" equilibria, this is not really a new equilibrium concept but rather a way of describing the equilibria of a particular class of games.)

If the players can condition their strategies on other variables in addition to calendar time, they may prefer not to use open-loop strategies in order to react to exogenous moves by nature, to the realizations of mixed strategies by their rivals, and to possible deviations by their rivals from the equilibrium strategies. That is, they may prefer to use closed-loop strategies. When closed-loop strategies are feasible, subgame-perfect equilibria will typically not be in open-loop strategies, as subgame perfection requires players to respond optimally to the realizations of random variables as well as to unexpected deviations; in particular, for open-loop strategies to meet this condition requires that it be optimal to play the same actions whether or not an opponent has deviated in the past. The term *closed-loop equilibrium* usually means a subgame-perfect equilibrium of the game where players can observe and respond to their opponents' actions at the end of each period. Of course, games with this information structure can have Nash equilibria that are not perfect. In particular, a pure-strategy open-loop equilibrium is a Nash equilibrium in the closed-loop information structure if the game is deterministic (admits no moves by nature) and the players' action spaces depend only on time (exercise 4.10); yet it will typically not be perfect.

It is typically much easier to characterize the open-loop equilibria of a given situation than the closed-loop ones, in part because the closed-loop strategy space is so much larger. This tractability is one explanation for the use of open-loop equilibria in economic analyses. A second reason for interest in open-loop equilibria, discussed in the next subsection, is that they serve as a useful benchmark for discussing the effects of strategic incentives in the closed-loop information structure, i.e., the incentives to change current play so as to influence the future play of opponents. A third reason, discussed in subsection 4.7.3, is that the open-loop equilibria may be a good approximation to the closed-loop ones if there are very many "small" players. Intuitively, if players are small, an unexpected deviation by an opponent might have little influence on a player's optimal play.

4.7.2 A Two-Period Example

The use of open-loop equilibria as benchmarks for measuring strategic effects can be illustrated most easily in a game with continuous action spaces. Consider a two-player two-stage game where in the first stage players $i = 1, 2$ simultaneously choose actions $a_i \in A_i$, and in the second stage they simultaneously choose actions $b_i \in B_i$, where each of these action sets is an interval of real numbers. Suppose that the payoff functions u_i are differentiable and that each player's payoff is concave in his own actions.

An open-loop equilibrium is a time path (a^*, b^*) satisfying, for $i = 1, 2$,

$$a_i^* \text{ maximizes } u_i((a_i, a_{-i}^*), b^*)$$

and

$$b_i^* \text{ maximizes } u_i(a^*, (b_i, b_{-i}^*)).$$

Since payoffs are concave, an interior solution must satisfy the first-order conditions

$$\frac{\partial u_i}{\partial a_i} = 0 = \frac{\partial u_i}{\partial b_i}. \tag{4.1}$$

In a closed-loop equilibrium (supposing that one exists), the second-stage actions $b^*(a)$ after any first-period actions a are required to be a Nash equilibrium of the second-stage game. That is, for each $a = (a_1, a_2)$, $b_i^*(a)$ maximizes $u_i(a, b_i, b_{-i}^*(a))$. Moreover, the players recognize that the second-period actions will depend on the first-period ones according to the function b^* when choosing the first-period actions. Thus, the first-order condition for an optimal (interior) choice of a_i (assuming that $b^*(\cdot)$ is differentiable) is now

$$\frac{\partial u_i}{\partial a_i} + \frac{\partial u_i}{\partial b_{-i}} \frac{\partial b_{-i}^*}{\partial a_i} = 0. \tag{4.2}$$

Compared with the corresponding open-loop equation (4.1), there is now an extra term corresponding to player i's "strategic incentive" to alter a_i to influence b_{-i}. (The change in player i's utility due to the induced change in his own second-stage action is 0 by the "envelope theorem.") For example, if player i prefers decreases in b_{-i}, and $\partial b_{-i}^*/\partial a_i$ is negative at the open-loop equilibrium a^*, then player i's "strategic incentive," at least locally, is to increase a_i beyond a_i^*.

To make these observations more concrete, suppose that the actions are choices of outputs, as in Cournot competition, and there is "learning by doing" so that a firm's second-stage marginal cost is decreasing in its first-stage output. Then the second-period equilibrium, $b^*(a)$, is simply the Cournot equilibrium given the first-period costs. Since a firm's Cournot-equilibrium output level is increasing in its opponent's marginal cost (at least if the stability condition discussed in chapter 1 is satisfied), and since increasing a_i lowers firm i's second-stage costs, $\partial b^*_{-i}/\partial a_i$ is negative. Finally, in Cournot competition each firm prefers its opponent's output to be low. Thus the strategic incentives in this example favor additional investment in learning beyond what a firm would choose in an open-loop equilibrium.

As a final gloss on this point, note that if the second-period equilibrium actions are increasing in the first-period actions, and firms prefer their opponents' second-period actions to be low, then strategic incentives tend to reduce first-period actions from the open-loop levels. And note that by changing the sign of $\partial u_i/\partial b_{-i}$ we obtain two analogous cases.[10]

4.7.3 Open-Loop and Closed-Loop Equilibria in Games with Many Players†††

We remarked above that one defense of open-loop equilibria is that they may approximate the closed-loop ones if there are many small players. We now examine this intuition in a bit more detail. First consider the limit case where players are infinitesimal. That is, suppose that the game has a continuum of nonatomic individuals of each player type —a continuum of player 1s, a continuum of player 2s, and so on. (For concreteness, let the set of individuals be copies of the unit interval, endowed with Lebesgue measure.) Suppose further that each player i's payoff is independent of the actions of any subset of opponents with measure 0. Then if one individual player j deviates, and all players $k \neq i,j$ ignore j's deviation, it is clearly optimal for player i to ignore the deviation as well. Thus, the outcome of an open-loop equilibrium is subgame perfect.[11]

However, even in this nonatomic model, there can be equilibria in which all players do respond to a deviation by a single player, because the deviation shifts play from one continuation equilibrium to another one. This is easiest to see in a continuum-of-individuals version of the two-stage game in figure 4.2. Suppose that each player 1's payoff to strategy s_1 is the average of his payoffs against the distribution of strategies played by the continuum of player 2's payoff:

$$u_1 = \sum_{s_2} p(s_2) u_1(s_1, s_2),$$

10. Bulow, Geanakoplos, and Klemperer (1985) and Fudenberg and Tirole (1984, 1986) develop taxonomies along these lines and apply them to various problems in industrial organization.

11. The open-loop *strategies* are not perfect, as they ignore deviations by subsets of positive measure.

where $p(s_2)$ is the proportion of player 2's using strategy s_2; define player 2's payoff analogously as the average against the population distribution of s_1's. If no player can observe the action of any individual opponent, then each player's second-stage payoff is independent of how he plays in the first stage, and the efficient first-stage payoff (5, 5) cannot occur in equilibrium. However, if players do observe the play of each individual opponent, then the first-stage payoff (5, 5) can be enforced with the same strategies as in the two-player version of the game.

The key to enforcing (5, 5) in this example is for all players to respond to the deviation of a single opponent, even though this deviation does not directly influence their payoffs. Economists fairly often rule out such "atomic" or "irregular" equilibria by requiring that players cannot observe the actions of measure-zero subsets.[12] However, these atomic equilibria are not pathologies of the continuum-of-players model. Fudenberg and Levine (1988) give an example of a sequence of two-period finite-player games, each of which has one open-loop equilibrium and two closed-loop ones, and such that every second-period subgame has a unique equilibrium for every finite number of players. As the number of players grows to infinity, one of the closed-loop equilibria has the same limiting path as the open-loop equilibrium, while the other closed-loop equilibrium converges to an atomic equilibrium of the limit game. To obtain the intuitive conclusion that open-loop and closed-loop equilibria are close together in T-period games with a large finite number of players requires (strong) conditions on the first $T + 1$ partial derivatives of the payoff functions.

It may be that the intuition that the actions of small players should be ignored corresponds to a continuum-of-players model that is the limit of finite-players models with a noise term that is large enough to mask the actions of any individual player, yet vanishes as the number of players grows to infinity so that the limit game is deterministic. We have not seen formal results along this line.

4.8 Finite-Horizon and Infinite-Horizon Equilibria (technical)[††]

Since continuity at infinity (see section 4.2) implies that events after t (for t large) have little effect, one would expect that under this condition the sets of equilibria of finite-horizon and infinite-horizon versions of the "same game" would be closely related. This is indeed the case, but it is not true that all infinite-horizon equilibria are limits of equilibria of the corresponding finite-horizon game. (That is, there is typically a failure of lower hemi-continuity in passing to the infinite-horizon limit.) We have already seen an example of this in the repeated prisoner's dilemma of figure 4.1.

12. For a recent example see Gul, Sonnenschein, and Wilson 1986.

Radner (1980) observed that cooperation can be restored as an equilibrium of the finite-horizon game (with time averaging, i.e., $\delta = 1$) if one relaxes the assumption that players exactly maximize their payoffs.

Definition 4.3 Profile σ^* is an ε-*Nash equilibrium* if, for all players i and strategies σ_i,

$$u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*) - \varepsilon.$$

The profile is an ε-perfect equilibrium if no player can gain more than ε by deviating in any subgame.

In our opinion, the concept of ε-equilibrium is best viewed as a useful device to relate the structures of large finite and infinite horizons. Though it is sometimes proposed as a description of boundedly rational behavior, its rationality requirements are very close to those of Nash equilibrium. For instance, the players must have the correct beliefs about their opponents' strategies and must correctly compute the expected payoff to each action; for some unspecified reason, they may voluntarily sacrifice ε utils.[13] (However, ε-optimality might be a *necessary* condition for certain boundedly rational policies to "survive.")

Radner's use of ε-equilibria smoothes the finite-to-infinite-horizon limit when utilities are continuous at infinity. Fudenberg and Levine (1983) showed this, beginning with an infinite-horizon game G^∞, which is then "approximated" by a sequence of T-period "truncated" games G^T that are created by choosing an arbitrary strategy $\tilde{s}$ in G^∞, and specifying that play follows $\tilde{s}$ after T. (In a repeated game, the most natural truncated games to consider are those in which $\tilde{s}$ specifies the same strategy in every period, independent of the history; these truncations correspond to the finitely repeated version of the game. In more general multi-stage games, such constant strategies may not be feasible.) A strategy s^T for the game G^T specifies play in all periods up to and including T; play follows strategy $\tilde{s}$ in the subsequent periods. In an abuse of notation, we use the same notation s^T to denote both the strategy of a truncated game G^T and the corresponding strategy in G^∞; when we speak of truncated-game strategies converging we will view them as elements of G^∞. Given this embedding, the payoff functions of G^∞ induce payoff functions in G^T in the obvious way.

To characterize equilibria of G^∞ in terms of limits of strategy profiles in G^T, one must specify a topology on the space of strategies of G^∞. Recall that a behavior strategy σ_i for player i in G^∞ is a sequence $\sigma_i(\cdot \mid h^0)$, $\sigma_i(\cdot \mid h^1)$, etc. Fudenberg and Levine use a complicated metric to topologize these

13. Another difficulty with the concept as a descriptive model is that, although ε is small relative to total utility, it may be large relative to a given period's utility. Thus, for instance, cooperating in the last period of a finitely repeated prisoner's dilemma may entail a substantial cost in that period even though the cost is negligible overall.

sequences.[14] In games with a finite number of possible actions per period ("finite-action games"), their topology reduces to the product topology (also known as the topology of pointwise convergence), which is much easier to work with.[15]

Theorem 4.4 (Fudenberg and Levine 1983) Consider an infinite-horizon finite-action game G^∞ whose payoffs are continuous at infinity. Then

(i) σ^* is a subgame-perfect equilibrium of G^∞ if and only if it is the limit (in the product topology) of a sequence σ^T of ε_T-perfect equilibria of a sequence of truncated games G^T with $\varepsilon_T \to 0$. Moreover,

(ii) the set of subgame-perfect equilibria of G^∞ is nonempty and compact.

Remark To gain some intuition for the theorem, consider approximating the finite-horizon prisoner's dilemma of figure 4.1, with $\delta > \frac{1}{2}$, by truncated games in which players are required to defect in all periods after T. Although "always defect" is the only subgame-perfect equilibrium of these finite games, cooperation can occur in an ε-perfect equilibrium: If the opponent's strategy is to cooperate until defecting occurs and to defect thereafter, the best response is to cooperate until the last period T, and then to defect at T, which yields average utility

$$1 + \frac{\delta^T(1-\delta)}{1-\delta^{T+1}}.$$

Cooperating in every period yields utility 1, and the difference between this strategy and the optimum goes to 0 as $T \to \infty$. More generally, continuity at infinity implies that players lose very little (in *ex ante* payoff) by not optimizing at a far-distant horizon.

Proof First note that if $\sigma^n \to \sigma$ in the product topology, then the continuation payoffs $u(\sigma^n|h^t)$ under σ^n in the subgame starting with h^t converge to the payoffs $u(\sigma|h^t)$ under σ. To see this, recall that $\sigma^n \to \sigma$ implies that

$$\sigma^n(a_i^t\,|\,h^t) \to \sigma(a_i^t\,|\,h^t) \text{ for all } t,\, h^t,\, a_i^t.$$

Thus, conditional on h^t, the probability that a^t is played in period t and

14. Harris (1985) shows that the complicated metric used by Fudenberg and Levine can be replaced by a simple one that sets the distance between two strategy profiles equal to $1/k$, where k is the largest number such that the two profiles prescribe exactly the same actions in every period $t \le k$ for every history h^t. This topology allows Harris to dispense with an additional continuity requirement that Fudenberg and Levine required for these games. (They required that payoffs as a function of the infinite history h^∞ be continuous in the product topology, which implies continuity in each period's realized action.) Borgers (1989) shows that the *outcomes* of infinite-horizon pure-strategy equilibria coincide with the limits (in the product topology) of the outcomes of finite-horizon pure-strategy ε-equilibria.

15. A sequence σ^n converges to σ in the product topology (or topology of pointwise convergence) if and only if, for all i, all h^t, and all $a_i^t \in A_i(h^t)$,

$$\sigma_i^n(a_i^t\,|\,h^t) \to \sigma_i(a_i^t\,|\,h^t),$$

which implies that $\sigma^n(a|h^t) \to \sigma(a|h^t)$.

a^{t+1} is played in period $(t+1)$ is

$$\sigma^n(a^t|h^t)\sigma^n(a^{t+1}|h^t,a^t) \to \sigma(a^t|h^t)\sigma(a^{t+1}|h^t,a^t),$$

and the distribution over outcomes at each date $\tau > t$ converges pointwise to that generated by σ. Thus, for any $\varepsilon > 0$ and T there is an N such that for all $n > N$ the distribution of actions from period t through period $(t+T)$ generated by σ^n, conditional on h^t, is within ε of the distribution generated by σ. Since outcomes after $(t+T)$ are unimportant for large T, continuation payoffs under σ^n converge to those under σ.[16]

To prove (i), note that continuity at infinity says that there is a sequence $\eta_T \to 0$ such that events after period T matter no more than η_T for each player. If σ is a subgame-perfect equilibrium of G^∞, the projection σ^T of σ onto G^T must be a $2\eta_T$-perfect equilibrium of G^T: For any i, h^t, and $\tilde{\sigma}_i$,

$$u_i(\sigma_i,\sigma_{-i}|h^t) \geq u_i(\tilde{\sigma}_i,\sigma_{-i}|h^t),$$

so from continuity at infinity

$$u_i(\sigma_i^T,\sigma_{-i}^T|h^t) + \eta_T \geq u_i(\tilde{\sigma}_i^T,\sigma_{-i}^T|h^t) - \eta_T.$$

Hence, σ is the limit of $2\eta_T$-equilibria of G^T.

Conversely, suppose $\sigma^T \to \sigma$ is a sequence of ε_T-perfect equilibria of G^T with $\varepsilon_T \to 0$. Continuity at infinity implies that each σ^T is an $(\varepsilon_T + \eta_T)$-perfect equilibrium of G^∞. If σ is not subgame perfect, there must be a time t and a history h^t where some player i can gain at least $2\varepsilon > 0$ by playing some $\hat{\sigma}_i$ instead of σ_i against σ_{-i}. But since $\sigma^T \to \sigma$ and payoffs are continuous, for T sufficiently large player i could gain at least ε by playing $\hat{\sigma}_i$ instead of σ_i^T against σ_{-i}^T, which contradicts $\varepsilon_T \to 0$.

To prove (ii), note first that for fixed $\tilde{s}$ each G^T is a finite multi-stage game and hence has a subgame-perfect equilibrium σ^T. (See exercise 3.5 and chapter 8.) Because the space of infinite-horizon strategies is compact (this is Tychonov's theorem[17]), the sequence σ^T has an accumulation point; this accumulation point is a perfect equilibrium of G^∞ from (i). Because payoffs are continuous, a standard argument shows that the set of subgame-perfect equilibria is closed, and closed subsets of compact sets are compact. ■

Radner considered the repeated prisoner's dilemma with time averaging, and observed that "both cooperate" *is* an ε-perfect equilibrium outcome of the finitely repeated game, with the required $\varepsilon \to 0$ as the number of periods tends to infinity. That Radner obtained this result with time

16. With time averaging, payoffs are not continuous in the product topology because they are not continuous at infinity. Consider the one-player game where the feasible actions each period are 0 and 1 and the player's stage-game payoff equals his action. Then the sequence of strategies σ^n given by $\sigma^n(0|h^t) = 1$ for $t < n$ and $\sigma^n(1|h^t) = 1$ for $t \geq n$ converges to $\sigma(0|h^t) = 1$ for all t and h^t in the product topology, and the discounted payoffs converge to 0 as well, but under time averaging the payoff of each σ^n is 1.

17. See, for example, Munkres 1975.

averaging is somewhat misleading, as, in general, games with time averaging are not well behaved: There can be exact equilibrium payoffs of a finitely repeated stochastic game that are not even ε-equilibrium payoffs of the infinitely repeated version.[18]

Exercises

Exercise 4.1

(a)** Consider the following modification of the stationary symmetric war of attrition developed in subsection 4.5.2: $L(\hat{t}) = -(1 - \delta^{\hat{t}})/(1 - \delta)$, $F(\hat{t}) = L(\hat{t}) + \delta^{\hat{t}}v$, and $B(\hat{t}) = L(\hat{t}) + \delta^{\hat{t}}qv$, with $q \leq \frac{1}{2}$, which corresponds to the assumption that if both animals stop fighting simultaneously then each has probability q of winning the prize. Characterize the symmetric stationary equilibrium. Compute the limit as the time period shrinks and show it is independent of q.

(b)*** Characterize the entire set of perfect-equilibrium outcomes.

Exercise 4.2** Consider the following modification of the perfect-information preemption game developed in the text. Players now choose two times: a time s to do a feasibility study and a time t to build a plant. In order to build a plant, the player must have done a feasibility study in some $s < t$, with s and t required to be odd for player 2 and even for player 1. Doing a feasibility study costs ε in present value, where ε is small,

18. Here is a one-player game where the (unique) infinite-horizon equilibrium payoff is less than the limit of the payoffs of finite-horizon equilibria: The player must decide when to chop down a tree, which grows by one unit per period from an initial size of 0 at date 0. If the tree is chopped down at the beginning of period t, the player receives a flow payoff of 0 each period before t, a flow of 1 in periods t through $2t - 1$, and 0 thereafter. If the player discounts flows at rate δ, the player's strategy in the infinite-horizon game is to chop down the tree at the time t^* that maximizes

$$\delta^t(1 + \delta + \cdots + \delta^{t-1}) = \frac{\delta^t(1 - \delta^t)}{1 - \delta};$$

so t^* is such that δ^{t^*} is as close as possible to $\frac{1}{2}$. Note that t^* is independent of the (finite or infinite) horizon as long as the horizon is large enough.

However, if the player's utility is the average payoff, the optimal strategy with finite horizon T is to cut at the first time t where $t \geq T/2$, yielding an average payoff that converges to $\frac{1}{2}$, yet *no* policy in the infinite-horizon problem yields a strictly positive payoff. Here the problem is that the sequence of finite-horizon strategies "cut at $T/2$" converges (in the product topology) to the limit strategy "never cut," but the limit of the corresponding payoffs is not equal to the payoff of the limit strategy, so that the payoff is not a continuous function of the strategy. Sorin (1986) identifies a different way the finite-to-infinite-horizon limit can be badly behaved. In his example, the finite-horizon equilibria involve one player using a behavior strategy that assigns probability roughly t/T to an action that sends the game to an absorbing state, so that the state is reached with probability close to 1 with a long horizon T. These strategies once again have a limit that assigns probability 0 to stopping (reaching the absorbing state) in each period, and indeed the equilibrium payoffs of the infinite-horizon game do not include the finite-horizon limits. Lehrer and Sorin (1989) provide conditions for the finite-to-infinite-horizon limit to be well behaved in one-player games with time averaging.

but this cost is recouped except for lost interest payments if only that player builds a plant. Thus, the payoff is $-\varepsilon$ if a player does a feasibility study and never builds, $-1 - \varepsilon$ if a player does a feasibility study and both build, and $14 - (t - 7)^2 - \varepsilon(1 - \delta^{t-s})$ if a player does a study, builds, and his opponent does not build. Show that the equilibrium outcome that survives iterated conditional dominance is for player 1 to pay for a study in period 2 and wait until period 6 to build. Explain why player 1 is now able to postpone building, and why player 2 cannot preempt by doing a study at $s - 1$. Relate this to "ε-preemption." (This exercise was provided by R. Wilson.)

Exercise 4.3* Consider a variant of Rubinstein's infinite-horizon bargaining game where partitions are restricted to be integer multiples of 0.01, that is, x can be 0, 0.01, 0.02, ..., 0.99, or 1. Characterize the set of subgame-perfect equilibria for $\delta = \frac{1}{2}$ and for δ very close to 1.

Exercise 4.4* Consider the I-player version of Rubinstein's game, which Moulin (1986) attributes to Dutta and Gevers. At dates $1, I + 1, 2I + 1, \ldots$, player 1 offers a division $(x_1, \ldots, x_I)$ of the pie with $x_i \geq 0$ for all i, and $\sum_{i=1}^{I} x_i \leq 1$. At dates $2, I + 2, 2I + 2, \ldots$, player 2 offers a division, and so on. When player i offers a division, the other players simultaneously accept or veto the division. If all accept, the pie is divided; if at least one vetoes, player $i + 1$ (player 1 if $i = I$) offers a division in the following period. Assuming that the players have common discount factor δ, show that, for all i, player i offering division

$$\left(\frac{1}{1 + \cdots + \delta^{I-1}}, \frac{\delta}{1 + \cdots + \delta^{I-1}}, \ldots, \frac{\delta^{I-1}}{1 + \cdots \delta^{I-1}}\right)$$

for players $i, i + 1, \ldots, i - 1$ at each date $(kI + i)$ and the other players' accepting is a subgame-perfect equilibrium outcome.

Exercise 4.5* Solve Ståhl's finite-horizon bargaining problem for T even and then for T odd, and show that the outcomes of the two cases converge to a common limit as $T \to \infty$.

Exercise 4.6** Admati and Perry (1988) consider the following model of infinite-horizon, perfect-information joint investment in a public good: Players $i = 1, 2$ take turns making investments $x_i(t)$ in the project, which is "finished" at the first date T at which

$$\sum_{i=1,2} \sum_{t=0}^{T} x_i(t) \geq K.$$

Players receive no benefits from the project until it is completed; if it is completed at date T, player i receives benefit $\delta_i^T V$. Players have a convex cost of investment $c_i(x_i)$, with $c_i(0) = 0$; thus, player i's total payoff is

$$\delta_i^T V - \sum_{t=0}^{T} \delta_i^t c_i(x_i(t)).$$

Use iterated conditional dominance to show that the game has a unique subgame-perfect equilibrium. Hint: First show that there is an $\bar{x}_1$ such that, if the investment to date exceeds $K - x_1$, it is conditionally dominant for the player on move to finish the project. Then argue that the second round of conditional dominance implies that there is an $\bar{x}_2$ such that, if investment to date $K(t)$ exceeds $K - x_1 - \bar{x}_2$ but does not exceed $K - \bar{x}_1$, the player on move should not invest less than $K - K(t) - \bar{x}_1$.

Exercise 4.7** Prove that, in a game of perfect information, no subgame-perfect strategy profile is removed by iterated conditional dominance.

Exercise 4.8**

(a) Consider the two-person Rubinstein-Ståhl model of section 4.4. The two players bargain to divide a pie of size 1 and take turns making offers. The discount factor is δ. Introduce "outside options" in the following way: At each period, the player whose turn it is to make the offer makes the offer; the other player then has the choice among (1) accepting the offer, (2) exercising his outside option instead, and (3) continuing bargaining (making an offer the next period). Let x_0 denote the value of the outside option. Show that, if $x_0 \leq \delta/(1 + \delta)$, the outside option has no effect on the equilibrium path. Comment. What happens if $x_0 > \delta/(1 + \delta)$?

(b) Consider an alternative way of formalizing outside options in bargaining. Suppose that there is an "exogenous risk of breakdown" of renegotiation (Binmore et al. 1986). At each period t, assuming that bargaining has gone on up to date t, there is probability $(1 - x)$ that bargaining breaks down at the end of period t if the period-t offer is turned down. The players then get x_0 each. Show that the "outside opportunity" x_0 matters even if it is small, and compute the subgame-perfect equilibrium.

(c) In their study of supply assurance, Bolton and Whinston (1989) consider a situation in which the outside option is endogenous. Suppose that there are three players: two buyers ($i = 1, 2$) and a seller ($i = 3$). The seller has one indivisible unit of a good for sale. Each buyer has a unit demand. The seller's cost of departing from the unit is 0 (the unit is already produced). The buyers have valuations v_1 and v_2, respectively. Without loss of generality, assume that $v_1 \geq v_2$. Bolton and Whinston consider a generalization of the Rubinstein-Ståhl process. At dates $0, 2, \ldots, 2k, \ldots$, the seller makes offers; at dates $1, 3, \ldots, 2k + 1, \ldots$, the buyers make offers. Buyers' offers are prices at which they are willing to buy and among which the seller may choose. The seller can make an offer to only a single buyer, as she has only one unit for sale (alternatively one could consider a situation in which the seller organizes an auction in each even period). Consider a stationary equilibrium and show that, if parties have the same discount factor and as the time between offers tends to 0, the parties' perfect-equilibrium payoffs converge to $v_1/2$ for both the seller and buyer 1 and to 0 for buyer 2 if $v_1/2 > v_2$, and to v_2 for the seller, $v_1 - v_2$ for buyer 1, and 0 for buyer 2 if $v_1/2 < v_2$. (For a uniqueness result see Bolton and Whinston 1989.)

Exercise 4.9** As shown by Rubinstein (see section 4.4 above), the alternating-move bargaining process between two players has a unique equilibrium. Shaked has pointed out that with $I \geq 3$ players there are many (subgame-) perfect equilibria (see Herrero 1985 for more details). Prove that with $I = 3$ players, and for discount factor $\delta > \frac{1}{2}$, any partition of the pie is the outcome of a perfect equilibrium.

The game is as follows: Three players bargain over the division of a pie of size 1. A division is a triple (x_1, x_2, x_3) of shares for each player, where $x_i \geq 0$, $\sum_{i=1}^{3} x_i = 1$. At dates $3k + 1, k = 0, 1, \ldots$, player 1 offers a division; if players 2 and 3 both accept, the game is over. If one or both of them veto, bargaining goes on. Similarly, at dates $3k + 2$ (respectively, $3k$), player 2 (respectively, player 3) makes the offer. The game stops once an offer by one player has been accepted by the other two players.

Show that, if $\delta > \frac{1}{2}$, any partition can be supported as a (subgame-) perfect equilibrium.

Exercise 4.10*

(a) Show that a pure-strategy open-loop Nash equilibrium of a deterministic game in which action spaces depend only on time is a closed-loop Nash equilibrium. Hint: Use the analogy with a control (single decision maker) problem.

(b) Does this result hold when the open-loop Nash equilibrium is in mixed strategies? When the players learn stochastic moves by nature?

References

Admati, A., and M. Perry. 1988. Joint projects without commitment. Mimeo.

Aumann, R. 1964. Mixed vs. behavior strategies in infinite extensive games. *Annals of Mathematics Studies* 52: 627–630.

Binmore, K. 1981. Nash bargaining theory 1–3. London School of Economics Discussion Paper.

Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. The Nash bargaining solution in economic modeling. *Rand Journal of Economics* 17: 176–188.

Bolton, P., and M. Whinston. 1989. Incomplete contracts, vertical integration, and supply assurance. Mimeo, Harvard University.

Borgers, T. 1989. Perfect equilibrium histories of finite and infinite horizon games. *Journal of Economic Theory* 47: 218–227.

Bulow, J., J. Geanakoplos, and P. Klemperer. 1985. Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy* 93: 488–511.

Dasgupta, P., and J. Stiglitz. 1980. Uncertainty, industrial structure, and the speed of R&D. *Bell Journal of Economics* 11: 1–28.

Fudenberg, D., R. Gilbert, J. Stiglitz, and J. Tirole. 1983. Preemption, leapfrogging, and competition in patent races. *European Economic Review* 22: 3–31.

Fudenberg, D., and D. Levine. 1983. Subgame-perfect equilibria of finite and infinite horizon games. *Journal of Economic Theory* 31: 227–256.

Fudenberg, D., and D. Levine. 1988. Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory* 44: 1–18.

Fudenberg, D., and J. Tirole. 1984. The fat cat effect, the puppy dog ploy and the lean and hungry look. *American Economic Review, Papers and Proceedings* 74: 361–368.

Fudenberg, D., and J. Tirole. 1985. Preemption and rent equalization in the adoption of new technology. *Review of Economic Studies* 52: 383–402.

Fudenberg, D., and J. Tirole. 1986. *Dynamic Models of Oligopoly*. Harwood.

Ghemawat, P., and B. Nalebuff. 1985. Exit. *Rand Journal of Economics* 16: 184–194.

Gilbert, R., and R. Harris. 1984. Competition with lumpy investment. *Rand Journal of Economics* 15: 197–212.

Gul, F., H. Sonnenschein, and R. Wilson. 1986. Foundations of dynamic monopoly and the Coase conjecture. *Journal of Economic Theory* 39: 155–190.

Harris, C. 1985. A characterization of the perfect equilibria of infinite horizon games. *Journal of Economic Theory* 37: 99–127.

Harris, C., and C. Vickers. 1985. Perfect equilibrium in a model of a race. *Review of Economic Studies* 52: 193–209.

Hendricks, K., A. Weiss, and C. Wilson. 1988. The war of attrition in continuous time with complete information. *International Economic Review* 29: 663–680.

Hendricks, K., and C. Wilson. 1989. The war of attrition in discrete time. Mimeo, State University of New York, Stony Brook.

Herrero, M. 1985. A strategic bargaining approach to market institutions. Ph.D. thesis, London School of Economics.

Katz, M., and C. Shapiro. 1986. How to license intangible property. *Quarterly Journal of Economics* 101: 567–589.

Kuhn, H. 1953. Extensive games and the problem of information. *Annals of Mathematics Studies* 28. Princeton University Press.

Lehrer, E., and S. Sorin. 1989. A uniform Tauberian theorem in dynamic programming. Mimeo.

Maynard Smith, J. 1974. The theory of games and evolution in animal conflicts. *Journal of Theoretical Biology* 47: 209–221.

Moulin, H. 1986. *Game Theory for the Social Sciences*. New York University Press.

Munkres, I. 1975. *Topology: A First Course*. Prentice-Hall.

Pearce, D. 1984. Rationalizable strategic behavior and the problem of perfection. *Econometrica* 52: 1029–1050.

Radner, R. 1980. Collusive behavior in non-cooperative epsilon equilibria of oligopolies with long but finite lives. *Journal of Economic Theory* 22: 121–157.

Reinganum, J. 1981a. On the diffusion of a new technology: A game-theoretic approach. *Review of Economic Studies* 48: 395–405.

Reinganum, J. 1981b. Market structure and the diffusion of new technology. *Bell Journal of Economics* 12: 618–624.

Rubinstein, A. 1982. Perfect equilibrium in a bargaining model. *Econometrica* 50: 97–110.

Shaked, A., and J. Sutton. 1984. Involuntary unemployment as a perfect equilibrium in a bargaining game. *Econometrica* 52: 1351–1364.

Simon, L. 1988. Simple timing games. Mimeo, University of California, Berkeley.

Sorin, S. 1986. Asymptotic properties of a non zero-sum stochastic game. Mimeo, Université de Strasbourg.

Stahl, I. 1972. *Bargaining Theory*. Stockholm School of Economics.

Whinston, M. 1986. Exit with multiplant firms. HIER DP 1299, Harvard University.

5 Repeated Games

The best-understood class of dynamic games is that of repeated games, in which players face the same "stage game" or "constituent game" in every period, and the player's overall payoff is a weighted average of the payoffs in each stage. If the players' actions are observed at the end of each period, it becomes possible for players to condition their play on the past play of their opponents, which can lead to equilibrium outcomes that do not arise when the game is played only once. One example of this in the repeated prisoner's dilemma of section 4.3 is the "unrelenting" strategy "cooperate until the opponent defects; if ever the opponent defects, then defect in every subsequent period." The profile where both players use this unrelenting strategy is a subgame-perfect equilibrium of the infinitely repeated game if the discount factor is sufficiently close to 1: even though each player could do better in the short run by defecting instead of cooperating, for a patient player this short-run gain is outweighed by the prospect of unrelenting future "punishment." Section 4.3 considers this equilibrium as well as the one where players defect each period; there are other equilibria as well. Our goal in this chapter is to present a more systematic treatment of general repeated games. (Surveys of the literature on repeated games have been published by Aumann (1986, 1989), Mertens (1987), and Sorin (1988). Mertens, Sorin, and Zamir (1990) give a detailed exposition of repeated games, with emphasis on "large" action spaces, and discuss the related topic of stochastic games.)

Because repeated games do not allow for past play to influence the feasible actions or payoff functions in the current period, they cannot be used to model such important phenomena as investment in productive machinery and learning about the physical environment. Nevertheless, repeated games may be a good approximation of some long-term relationships in economics and political science—particularly those where "trust" and "social pressure" are important, such as when informal agreements are used to enforce mutually beneficial trades without legally enforced contracts. There are many variations on this theme, including Chamberlin's (1929) informal argument that oligopolists may use repeated play to implicitly collude on higher prices[1] and Macaulay's (1963) observation that relations between a firm and its suppliers are often based on "reputation" and the threat of the loss of future business.[2] Chapter 9 discusses an alternative way of modeling long-run relationships, where past actions serve to signal a player's future intentions by providing information about his payoffs.

1. Fisher (1898) gave an earlier critique of the static Cournot model that can be interpreted as favoring a repeated-game model. He asserted that, contrary to the Cournot assumption that outputs are chosen once and for all, "no business man assumes ... that his opponents' output or price will remain constant" (quoted in Scherer 1980).

2. Recent economic applications of repeated games to explain trust and cooperation include Greif 1989, Milgrom, North, and Weingast 1989, Porter 1983a, and Rotemberg and Saloner 1986. For some recent applications to political science, see the essays in Oye 1986.

The reason repeated play introduces new equilibrium outcomes is that players can condition their play on the information they have received in previous stages. Thus, one would expect that a key issue in analyzing repeated games is just what form this information takes. In the prisoner's-dilemma example in chapter 4, the players perfectly observed the actions that had been played. Sections 5.1–5.4 discuss general repeated games with this information structure, which we call *repeated games with observed actions*. (Note that this is a special case of the multi-stage games with observed actions introduced in chapter 3.)

Section 5.1 analyzes the equilibria of infinite-horizon games, focusing on the "folk theorems," which describe the equilibria when players are either completely patient or almost so. Section 5.2 presents the parallel results for finitely repeated games, and section 5.3 discusses various extensions to models where not all the players play the game every period. Examples include a long-run firm that faces a different short-run consumer each period; in that case the firm must decide whether to produce high-quality or low-quality goods (Dybvig and Spatt 1980; Shapiro 1982), and an organization composed of overlapping generations of workers must decide whether to exert effort on joint production (Crémer 1986).

Section 5.4 discusses the ideas of Pareto perfection and renegotiation-proofness, which have been proposed as a way to restrict the large set of repeated game equilibria when players are patient.

Sections 5.5–5.7 consider repeated games in which the players observe imperfect signals of their opponents' play. One example of this sort of game is the oligopoly model of Green and Porter (1984), wherein firms choose quantities each period and observe the realized market price but not the outputs of their opponents. Since the market price is stochastic, a low price could be due either to unexpectedly low demand or to some rival's having produced an unexpectedly high output. A second example is the repeated partnership in which each player observes the realized level of production but not the effort level of his partner (Radner 1986; Radner, Myerson, and Maskin 1986).

5.1 Repeated Games with Observable Actions††

5.1.1 The Model

The building block of a repeated game, the game which is repeated, is called the *stage game*. Assume that the stage game is a finite I-player simultaneous-move game with finite action spaces A_i and stage-game payoff functions $g_i\colon A \to \mathbb{R}$, where $A = \times_{i \in \mathcal{I}} A_i$. Let $\mathcal{A}_i$ be the space of probability distributions over A_i.

To define the repeated game, we must specify the players' strategy spaces and payoff functions. This section considers games in which the players observe the realized actions at the end of each period. Thus, let $a^t \equiv (a_1^t, \ldots, a_I^t)$ be the actions that are played in period t. Suppose that the game begins in period 0, with the null history h^0. For $t \geq 1$, let $h^t = (a^0, a^1, \ldots, a^{t-1})$ be the realized choices of actions at all periods before t, and let $H^t = (A)^t$ be the space of all possible period-t histories.

Since all players observe h^t, a pure strategy s_i for player i in the repeated game is a sequence of maps s_i^t—one for each period t that map possible period-t histories $h^t \in H^t$ to actions $a_i \in A_i$. (Remember that a strategy must specify play in all contingencies, even those that are not expected to occur.) A mixed (behavior) strategy σ_i in the repeated game is a sequence of maps σ_i^t from H^t to mixed actions $\alpha_i \in \mathscr{A}_i$. Note that a player's strategy cannot depend on the past values of his opponents' randomizing probabilities α_{-i}; it can depend only on the past values of a_{-i}. Note also that each period of play begins a proper subgame. Moreover, since moves are simultaneous in the stage game, these are the only proper subgames, a fact that we will use in testing for subgame perfection.[3]

This section considers infinitely repeated games; section 5.2 considers games with a fixed finite horizon. With a finite horizon, the set of subgame-perfect equilibria is determined by backward-induction arguments that do not apply to the infinite-horizon model. The infinite-horizon case is a better description of situations where the players always think the game extends one more period with high probability; the finite-horizon model describes a situation where the terminal date is well and commonly foreseen.[4]

There are several alternative specifications of payoff functions for the infinitely repeated game. We will focus on the case where players discount future utilities using discount factor $\delta < 1$. In this game, denoted $G(\delta)$, player i's objective function is to maximize the normalized sum

$$u_i = \mathrm{E}_\sigma(1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i(\sigma^t(h^t)),$$

where the operator E_σ denotes the expectation with respect to the distribution over infinite histories that is generated by strategy profile σ. The normalization factor $(1 - \delta)$ serves to measure the stage-game and repeated-game payoffs in the same units: The normalized value of 1 util per period is 1.

3. Although it seems that many of the results in this chapter should extend to stage games in which the moves are not simultaneous, as far as we know no one has yet checked the details.

4. The importance of a common forecast of the terminal date is shown by Neyman (1989), who considers a repeated prisoner's dilemma where both players know the horizon is finite, and where both players know the true length of the game to within ± 1 period, but the length of the game is not common knowledge between them (it is not even "almost common knowledge," as defined in chapter 14 below). He shows that this game has "cooperative" equilibria of the sort that arise in the infinite-horizon model but that are ruled out by backward induction with a known finite horizon.

To recapitulate the notation: As in the rest of the book, u_i, s_i, and σ_i denote the payoffs and the pure and mixed strategies of the overall game. The payoffs and strategies of the stage game are denoted g_i, a_i, and α_i.

As in the games of chapter 4, the discount factor δ can be thought of as representing pure time preference: This interpretation corresponds to $\delta = e^{-r\Delta}$, where r is the rate of time preference and Δ is the length of the period. The discount factor can also represent the possibility that the game may terminate at the end of each period: Suppose that the rate of time preference is r, the period length is Δ, and there is probability μ of continuing from one period to the next. Then 1 util tomorrow, to be collected only if the game lasts that long, is worth nothing with probability $1 - \mu$ and worth $\delta = e^{-r\Delta}$ utils with probability μ, for an expected discounted value of $\delta' = \mu\delta$. Thus, the situation is the same as if $\mu' = 1$ and $r' = r - \ln(\mu)/\Delta$. This shows that infinitely repeated games can represent games that terminate in finite time with probability 1. The key is that the conditional probability of continuing one more period should be bounded away from 0.[5]

Since each period begins a proper subgame, for any strategy profile σ and history h^t we can compute the players' expected payoffs from period t on. We will call these the "continuation payoffs," and renormalize so that the continuation payoffs from time t are measured in time-t units. Thus, the continuation payoff from time t on is

$$(1 - \delta) \sum_{\tau = t}^{\infty} \delta^{\tau - t} g_i(\sigma^\tau(h^\tau)).$$

With this renormalization, the continuation payoff of a player who will receive 1 util per period from period t on is 1 unit for any period t. This renormalization will be convenient, as it exploits the stationary structure of the game.

Although we will focus on the case where players discount future payoffs, we will also discuss the case where players are "completely patient," corresponding to the limit model $\delta = 1$. Several different specifications of the payoffs have been proposed to model complete patience. The simplest is the time-average criterion, where each player i's objective is to maximize

$$\liminf_{T \to \infty} \mathrm{E}\,(1/T) \sum_{t=0}^{T} g_i(\sigma^t(h^t)).$$

The lim inf in this expression is in response to the fact that some infinite sequences of utilities do not have well-defined average values.[6] (See Lehrer

5. In unpublished notes, B. D. Bernheim has shown that if the stage game has a continuum of actions, cooperative equilibria can arise even if the continuation probability does converge to 0 over time, provided it does so sufficiently slowly.

1988 for a discussion of the difference between this notion of a time average and the analogous one using the lim sup.)

Any form of time-average criterion implies that players are unconcerned not only about the timing of payoffs but also about their payoff in any finite number of periods, so that, for example, the sequences $(1, 0, 0, \ldots)$ and $(0, 0, \ldots)$, which both have average 0, are equally attractive. The overtaking criterion is an alternative specification of "patience" where improvement in a single period matters. This criterion, which is not representable by a utility functional, says that the sequence $g = (g^0, g^1, \ldots)$ is preferred to $\tilde{g} = (\tilde{g}^0, \tilde{g}^1, \ldots)$ if and only if there exists a time T' such that for all $T > T'$ the partial sum $\sum_{t=0}^{T} g^t$ strictly exceeds the partial sum $\sum_{t=0}^{T} \tilde{g}^t$. If g is not preferred to $\tilde{g}$, and $\tilde{g}$ is not preferred to g, then the two sequences are judged to be equally attractive. Note that if g has a higher time average than $\tilde{g}$, then g is necessarily preferred to $\tilde{g}$ under the overtaking criterion.[7]

Now that we have specified strategy spaces and payoff functions for the repeated game, our description of the model is complete. We conclude this subsection with a simple but useful observation.

Observation If α^* is a Nash equilibrium of the stage game (that is, a "static equilibrium"), then the strategies "each player i plays α_i^* from now on" are a subgame-perfect equilibrium. Moreover, if the game has m static equilibria $\{\alpha^j\}_{j=1}^m$, then for any map $j(t)$ from time periods to indices the strategies "play $\alpha^{j(t)}$ in period t" are a subgame-perfect equilibrium as well.

To see that this observation is correct, note that with these strategies the future play of player i's opponents is independent of how he plays today, so his optimal response is to play to maximize his current period's payoff, i.e., to play a static best response to $\alpha_{-i}^{j(t)}$. Note also that these are "open-loop" strategies of the type discussed in section 4.7.

The observation shows that repeated play of a game does not decrease the set of equilibrium payoffs. Further, since the only reason not to play a static best response is concern about the future, if the discount factor is small enough, then the only Nash equilibria of the repeated game are strategies that specify a static equilibrium at every history to which the equilibrium gives positive probability. (Proving this is exercise 5.2. Note that the same static equilibrium need not occur in every period. In games with infinite strategy spaces the conclusion must be modified slightly, since

6. Recall that $\lim_{t \to \infty} \inf x^t = \sup_T \inf_{t > T} x^t$ is the greatest lower bound on the sequence's accumulation points. Thus, if $\liminf_{t \to \infty} x^t = \underline{x}$, then for all $x > \underline{x}$ and all T there is a $\tau > T$ with $x^\tau < x$.

7. It is not obvious how to extend the overtaking criterion to probability distributions over sequences. One formulation requires that with probability 1 the realized sequence of utilities under one distribution be preferred to that under the other, but with this formulation the overtaking criterion is no longer a refinement of time averaging.

even a small future punishment can induce players to forgo a sufficiently small current gain.)

Another important fact about repeated games with observed actions is that the set of Nash-equilibrium continuation payoff vectors is the same in every subgame. Proving this is exercise 5.3.

5.1.2 The Folk Theorem for Infinitely Repeated Games

The "folk theorems" for repeated games assert that if the players are sufficiently patient then any feasible, individually rational payoffs can be enforced by an equilibrium. Thus, in the limit of extreme patience, repeated play allows virtually any payoff to be an equilibrium outcome.

To make this assertion precise, we must define "feasible" and "individually rational." Define player i's *reservation utility* or *minmax value* to be

$$v_i = \min_{\alpha_{-i}} \left[\max_{\alpha_i} g_i(\alpha_i, \alpha_{-i}) \right]. \tag{5.1}$$

This is the lowest payoff player i's opponents can hold him to by any choice of α_{-i}, provided that player i correctly foresees α_{-i} and plays a best response to it. Let m^i_{-i} be a strategy for player i's opponents that attains the minimum in equation 5.1. We call m^i_{-i} the *minmax profile* against player i. Let m^i_i be a strategy for player i such that $g_i(m^i_i, m^i_{-i}) = v_i$.

To illustrate this definition, we compute the minmax values for the game in figure 5.1. To compute player 1's minmax value, we first compute his payoffs to U, M, and D as a function of the probability q that player 2 assigns to L; in the obvious notation, these payoffs are $v_U(q) = -3q + 1$, $v_M(q) = 3q - 2$, and $v_D(q) = 0$. Since player 1 can always attain a payoff of 0 by playing D, his minmax payoff is at least this large; the question is whether player 2 can hold player 1's maximized payoff to 0 by some choice of q. Since q does not enter into v_D, we can pick q to minimize the maximum of v_U and v_M, which occurs at the point where the two expressions are equal, i.e., $q = \frac{1}{2}$. Since $v_U(\frac{1}{2}) = v_M(\frac{1}{2}) = -\frac{1}{2}$, player 1's minmax value is the zero payoff he can achieve by playing D. (Note that $\max(v_U(q), v_M(q)) \leq 0$ for any $q \in [\frac{1}{3}, \frac{2}{3}]$, so we can take player 2's minmax strategy against player 1, m^1_2, to be any q in this range.)

Similarly, to find player 2's minmax value, we first express player 2's payoff to L and R as a function of the probabilities p_U and p_M that player 1 assigns to U and M:

	L	R
U	−2,2	1,−2
M	1,−2	−2,2
D	0,1	0,1

Figure 5.1

$$v_L = 2(p_U - p_M) + (1 - p_U - p_M), \tag{5.2}$$

$$v_R = -2(p_U - p_M) + (1 - p_U - p_M). \tag{5.3}$$

Player 2's minmax payoff is then determined by

$$\min_{p_U, p_M} \max[2(p_U - p_M) + (1 - p_U - p_M),$$
$$- 2(p_U - p_M) + (1 - p_U - p_M)].$$

By inspection (or plotting equations 5.2 and 5.3) we see that player 2's minmax payoff is 0, which is attained by the profile $(\frac{1}{2}, \frac{1}{2}, 0)$. Here, unlike the minmax against player 1, the minmax profile is uniquely determined: If $p_U > p_M$, the payoff to L is positive, if $p_M > p_U$ the payoff to R is positive, and if $p_U = p_M < \frac{1}{2}$, then both L and R have positive payoffs.

Note that if we restricted attention to pure strategies in equation 5.1, player 1's and player 2's minmax values will both be 1. Clearly, minimizing over a smaller set in equation 5.1 cannot give a lower value; the figure shows that the restriction can give values that are strictly higher.

At this point, the reader might question our identifying the minmax payoffs as the reservation utilities. This terminology is justified by the following observation.

Observation Player i's payoff is at least $\underline{v}_i$ in any static equilibrium and in any Nash equilibrium of the repeated game, regardless of the level of the discount factor.

Proof In a static equilibrium $\hat{\alpha}$, $\hat{\alpha}_i$ is a best response to $\hat{\alpha}_{-i}$, and so $g_i(\hat{\alpha}_i, \hat{\alpha}_{-i})$ is no less than the minimum defined in equation 5.1. Now consider a Nash equilibrium $\hat{\sigma}$ of the repeated game. One feasible, though not necessarily optimal, strategy for player i is the myopic one that chooses each period's action $a_i(h^t)$ to maximize the expected value of $g_i(a_i, \hat{\sigma}_{-i}(h^t))$. (This may not be optimal, because it ignores the possibility that the future play of i's opponents may depend on how he plays today.) The key is that, because all players have the same information at the start of each period t, the probability distribution over the opponents' period-t actions given player i's information corresponds to *independent* randomizations by player i's opponents. (This is not necessarily true when actions are imperfectly observed, as we discuss in section 5.5.) Thus, the myopic strategy for player i yields at least $\underline{v}_i$ in each period, and so $\underline{v}_i$ is a lower bound on player i's equilibrium payoff in the repeated game. ∎

Thus, we know on *a priori* grounds that no equilibrium of the repeated game can give any player a payoff lower than his minmax value.

Next we introduce a definition of the feasible payoffs. Here we encounter the following subtlety: The sets of feasible payoffs in the stage game, and thus in the repeated game for small discount factors, need not be convex.

The problem is that "many" convex combinations of pure-strategy payoffs correspond to correlated strategies, and cannot be obtained by independent randomizations. For example, in the "battle of the sexes" game (figure 1.10a), the payoffs $(\frac{3}{2}, \frac{3}{2})$ cannot be obtained by independent mixing.

As Sorin (1986) has shown, this nonconvexity does not occur when the discount factor is near enough to 1, as any convex combination of pure-strategy payoffs can be obtained by a time-varying deterministic path. This is easiest to see in the time-averaging limit: The payoffs $(\frac{3}{2}, \frac{3}{2})$ in the battle of the sexes can be obtained by playing (B, B) in even-numbered periods and (F, F) in odd-numbered ones.

To avoid the need to use such time-varying paths, Fudenberg and Maskin (1986a) convexify the feasible payoffs of the stage game by assuming that all players observe the outcome of a public randomizing device at the start of each period. Sorin's result on its own suggests, but does not imply, that these public randomizations are innocuous when the discount factor is near enough to 1; Fudenberg and Maskin (1990a) subsequently proved a stronger version of Sorin's result and used it to extend their proof of the perfect folk theorem to games without public randomizations.[8] To avoid the complications this involves, we will use the assumption of public randomizations in our proofs. Formally, let $\{\omega^0, \ldots, \omega^t \ldots\}$ be a sequence of independent draws from a uniform distribution on [0, 1], and assume that the players observe ω^t at the beginning of period t. The history is now

$$h^t \equiv (a^0, \ldots, a^{t-1}, \omega^0, \ldots, \omega^t).$$

A pure strategy s_i for player i is then a sequence of maps s_i^t from histories h^t into A_i.

In this case, the set of feasible payoffs for any discount factor is

$$V = \text{convex hull}\{v \mid \exists a \in \mathrm{A} \text{ with } g(a) = v\}.$$

This set is illustrated in figure 5.2. The shaded region in the figure is the set of all feasible payoffs that Pareto dominate the minmax payoffs, which are 0 for both players. The set of feasible, strictly individually rational payoffs is the set $\{v \in V \mid v_i > \underline{v}_i \ \forall i\}$. Figure 5.2 depicts these sets for the game of figure 5.1, in which the minmax payoffs are (0, 0).

Theorem 5.1 (folk theorem)[9] For every feasible payoff vector v with $v_i > \underline{v}_i$ for all players i, there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a Nash equilibrium of $G(\delta)$ with payoffs v.

8. For small discount factors public randomizations can allow equilibrium payoffs that are not in the convex hull of the set of equilibrium payoffs without public randomization. See Forges 1986, Myerson 1986, and our exercise 5.5.

9. This is called the "folk theorem" because it was part of game theory's oral tradition or "folk wisdom" long before it was recorded in print.

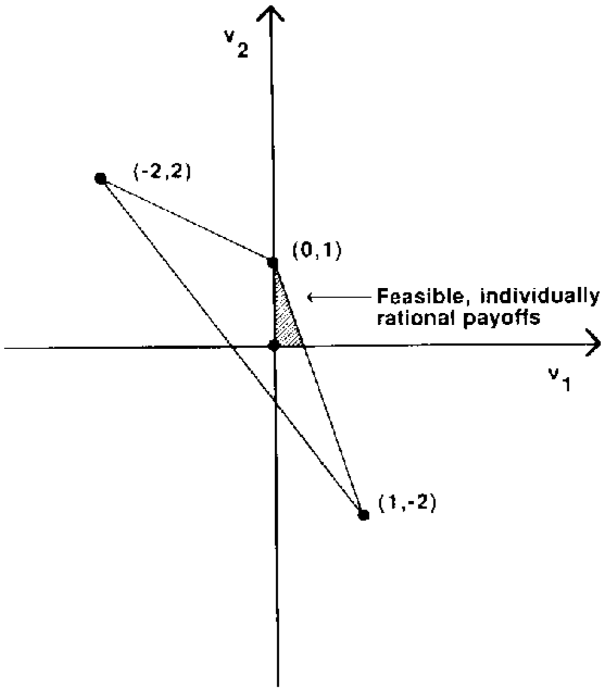

Figure 5.2

Remark The intuition for this theorem is simply that when the players are patient, any finite one-period gain from deviation is outweighed by even a small loss in utility in every future period. The strategies we construct in the proof are "unrelenting": A player who deviates will be minmaxed in every subsequent period.

Proof Assume first that there is a pure action profile a such that $g(a) = v$, and consider the following strategies for each player i: "Play a_i in period 0, and continue to play a_i so long as either (i) the realized action in the previous period was a or (ii) the realized action in the previous period differed from a in two or more components. If in some previous period player i was the only one not to follow profile a, then each player j plays m_j^i for the rest of the game."

Can player i gain by deviating from this strategy profile? In the period in which he deviates he receives at most $\max_a g_i(a)$, and since his opponents will minmax him forever afterward he can obtain at most $\underline{v}_i$ in periods after his first deviation. Thus, if player i's first deviation is in period t, he obtains at most

$$(1 - \delta^t)v_i + \delta^t(1 - \delta) \max_a g_i(a) + \delta^{t+1}\underline{v}_i, \tag{5.4}$$

which is less than v_i as long as δ exceeds the critical level $\underline{\delta}_i$ defined by

$$(1 - \underline{\delta}_i) \max_a g_i(a) + \underline{\delta}_i \underline{v}_i = v_i. \tag{5.5}$$

Since $v_i > \underline{v}_i$, the solution $\underline{\delta}_i$ for equation 5.5 is less than 1. Taking $\underline{\delta} = \max_i \underline{\delta}_i$ completes the argument. Note that, in deciding whether to deviate, in period t, player i assigns probability 0 to an opponent deviating in the same period. This is a consequence of the definition of Nash equilibrium: Only unilateral deviations are considered.

If payoffs v cannot be generated using pure actions, then we replace the action profile a with a public randomization $a(\omega)$ yielding payoffs with expected value v. The discount factor required to ensure that player i cannot gain by deviating may be somewhat larger in this case, as if player i conforms he does not receive exactly v_i in each period, and his temptation to deviate may be greater in periods where $g_i(a(\omega))$ is relatively low. It will be sufficient to take $\underline{\delta}_i$ such that

$$(1 - \delta_i) \max_a g_i(a) + \underline{\delta}_i v_i = (1 - \underline{\delta}_i) \min_a g_i(a) + \underline{\delta}_i v_i. \tag{5.6}$$

To see that equation 5.6 is sufficient, note that for any period-t realization of ω, player i's continuation payoff to conforming from t on is

$$(1 - \delta) g_i(a(\omega)) + \delta v_i,$$

which is at least as large as $(1 - \delta)\min_a g_i(a) + \delta v_i$. By hypothesis, δ is large enough that this latter expression exceeds the continuation payoff from deviating, which is at most $(1 - \delta)\max_a g_i(a) + \delta \underline{v}_i$. ■

Under the strategies used in the proof of theorem 5.1, a single deviation provokes unrelenting punishment. Now, such punishments may be very costly for the punishers to carry out. For example, in a repeated quantity-setting oligopoly, the minmax strategies require player i's opponents to produce so much output that price falls below player i's average cost, which may be below their own costs as well. Since minmax punishments can be costly, the question arises if player i ought to be deterred from a profitable one-shot deviation by the fear that his opponents will respond with the unrelenting punishment specified above. More formally, the point is that the strategies we used to prove the Nash folk theorems are not subgame perfect. This raises the question of whether the conclusion of the folk theorem applies to the payoffs of perfect equilibrium.

The answer to this question is yes, as shown by the perfect folk theorem. Friedman (1971) proved a weaker result, sometimes called a "Nash-threats" folk theorem.

Theorem 5.2 (Friedman 1971) Let α^* be a static equilibrium (an equilibrium of the stage game) with payoffs e. Then for any $v \in V$ with $v_i > e_i$ for all players i, there is a $\underline{\delta}$ such that for all $\delta > \underline{\delta}$ there is a subgame-perfect equilibrium of $G(\delta)$ with payoffs v.

Proof Assume that there is an $\hat{a}$ with $g(\hat{a}) = v$, and consider the following strategy profile: In period 0 each player i plays $\hat{a}_i$. Each player i continues to play $\hat{a}_i$ so long as the realized actions were $\hat{a}$ in all previous periods. If at least one player did not play according to $\hat{a}$, then each player i plays α_i^* for the rest of the game.

This strategy profile is a Nash equilibrium for δ large enough that

$$(1 - \delta) \max_a g_i(a) + \delta e_i < v_i. \tag{5.7}$$

This inequality is satisfied for a range of δ less than 1 because it holds strictly at the $\delta = 1$ limit. To check that the profile is subgame perfect, note that in every subgame off the equilibrium path the strategies are to play α^* forever, which is a Nash equilibrium for any static equilibrium α^*.

If there is no $\hat{a}$ with $g(\hat{a}) = v$, we use public randomizations as in the previous theorem. ∎

Friedman's result shows that patient, identical Cournot duopolists can "implicitly collude" by each producing half of the monopoly output, with any deviation triggering a switch to the Cournot outcome forever after. This equilibrium is "collusive" in obtaining the monopoly price; the collusion is "implicit" in that it can be enforced without the use of binding contracts. Instead, each firm is deterred from breaking the agreement by the (credible) fear of provoking Cournot competition.

There is ample evidence that firms in some industries have understood the role of repeated play in allowing such collusive outcomes (although other models than the repeated games considered here can be used to capture the effects of repeated play). Some of the agents involved have even recognized the key role of the interval between periods in determining whether the discount factor is large enough to allow collusion to be an equilibrium, and have suggested that the industry take steps to ensure that any defectors from the collusive outcome will be detected quickly. Scherer (1980) quotes the striking example of the American Hardwood Manufacturer's Association, which proclaimed: "Knowledge regarding prices actually made is all that is necessary to keep prices at reasonably stable and normal levels.... By keeping all members fully and quickly informed of what others have done, the work of the Plan results in a certain uniformity of trade practices.... Cooperative competition, not cut-throat competition."

The conclusion of Friedman's theorem is weaker than that of the folk theorem, except in games with a static equilibrium that holds all the players to their minmax values. (This is a fairly special condition, but it does hold in the prisoner's dilemma and in Bertrand competition with perfect substitutes and constant returns to scale.) Thus, Friedman's theorem leaves open the question of whether the requirement of perfect equilibrium restricts the limit set of equilibrium payoffs. The "perfect folk theorems" of Aumann and Shapley (1976), Rubinstein (1979a), and Fudenberg and Maskin (1986a) show that this is not the case: For any feasible, individually rational payoff vector, there is a range of discount factors for which that payoff vector can be obtained in a subgame-perfect equilibrium.

As a first step toward understanding the strategies used in the folk theorems, note that to hold player i's payoff very near to his minmax value in an equilibrium, his opponents' must specify that if player i deviates from the equilibrium path they will "punish" him by playing the minmax strategies m^i_{-i} (or a profile very close to it) for at least one period. (Otherwise, if player i were to play a static best response to his opponents' strategies in every period, his payoff in every period would be bounded away from his minmax value, and so his overall payoff would exceed his minmax value as well.) Thus, the perfect folk theorem requires that there be perfect-equilibrium strategies in which player i's opponents play m^i_{-i}. It is easy to induce player i's opponents to play m^i_{-i} for a finite number of periods when intertemporal preferences are represented by the time-average criterion, as then, even if punishment reduces the punishers' per-period payoff, the overall cost of the punishment is 0. This is the intuition for the following theorem.

Theorem 5.3 (Aumann and Shapley 1976) If players evaluate sequences of stage-game utilities by the time-average criterion, then for any $v \in V$ with $v_i > \underline{v}_i$ for all players i, there is a subgame-perfect equilibrium with payoffs v.

Proof Consider the following strategies: "Begin in the 'cooperative phase.' In this phase, play a public randomization ρ with payoff v, and remain in this phase as long as there are no deviations. If player i deviates, play the minmax strategy $m^i = (m^i_i, m^i_{-i})$ for N periods, where N is chosen so

$$\max_a g_i(a) + N\underline{v}_i < \min_a g_i(a) + Nv_i,$$

for all players i. After the N periods have elapsed, return to the cooperative phase, regardless of whether there were any deviations from m^i."

Recall that the one-stage deviation principle does not apply in infinite-horizon games with time averaging. Hence, to verify that these strategies are a perfect equilibrium, we must explicitly verify that there is no strategy that improves a player's payoff in any subgame. The condition on N ensures that any gains from deviation in the cooperative phase are removed at the punishment phase, so no sequence of a finite or infinite number of deviations can increase player i's average payoff above v_i. Moreover, even though minmaxing a deviator is costly in terms of per-period payoff, any finite number of such losses are costless with the time-average criterion. Thus, player j's average payoff in a subgame where player i is being punished is v_j, and no player j can gain by deviating in any subgame. Therefore the strategies are subgame perfect. ■

The strategies in this last proof are not subgame perfect under the overtaking criterion studied by Rubinstein (1979a), since with that criterion players do care about the finite number of periods where they may incur a loss by minmaxing an opponent. To prove the folk theorem for this case, Rubinstein used strategies in which the punishment lengths grow exponentially: The first deviator is punished for N periods, a player who deviates from minmaxing the first deviator is punished for N^2 periods, a player who deviates from punishing a player who deviated from punishing the first deviator is punished for N^3 periods, and so on. Here, N is chosen long enough so that it is better for any player to minmax any opponent for one period and then have play revert to payoffs v per period, than to be minmaxed for N periods, and then have per-period payoffs revert to v.

When the players discount their future payoffs, this kind of scheme will not work: If player i's payoff when minmaxing player j, $g_i(m^j)$, is strictly less than his own minmax value, $\underline{v}_i$, then for any $\delta < 1$ there is a k where punishing player j for N^k periods is not individually rational: The best possible payoff from playing m^j for N^k periods is

$$(1 - \delta^{N^k})g_i(m^j) + \delta^{N^k} \max_a g_i(a),$$

which converges to $g_i(m^j) < \underline{v}_i$ as k goes to infinity.

Thus, to obtain the folk theorem in the limit of discount factors tending to 1, Fudenberg and Maskin (1986a) consider a different type of strategy—one that induces player i's opponents to minmax him not by threatening them with "punishments" if they don't minmax but rather by offering them "rewards" if they do. Abreu (1986, 1988) makes the same observation in his work on the structure of the equilibrium set for fixed discount factors; we discuss his results below. Now, in designing strategy profiles that provide such rewards for punishing a deviator, one must take care not to reward the original deviator as well, for such a reward could undo the effect of the punishments and make deviations attractive. The need to be able to provide rewards for punishing player i without rewarding player i himself leads to the "full-dimension" condition used in the following theorem.

Theorem 5.4 (Fudenberg and Maskin 1986a) Assume that the dimension of the set V of feasible payoffs equals the number of players. Then, for any $v \in V$ with $v_i > \underline{v}_i$ for all i, there is a discount factor $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a subgame-perfect equilibrium of $G(\delta)$ with payoffs v.

Remarks

(1) Fudenberg and Maskin give an example of a three-player game with $\dim V = 1$ where the folk theorem fails. Abreu and Dutta (1990) weaken the full-dimension condition to $\dim V = I - 1$; Smith (1990) shows that it suffices that the projection of V^* onto the coordinate space of any two players' payoffs is two-dimensional.

(2) Rubinstein's version of the perfect folk theorem supposes that any deviation from a minmax profile was certain to be detected, which requires

either that the minmax profile be in pure actions or that the players' choices of randomizing probabilities, and not just their realized actions, are observed at the end of each period. As is noted above, the restriction to pure minmax strategies can lead to higher minmax values. Indeed, the pure-strategy minmax values can be above the payoffs in any static equilibrium.

Proof

(i) For simplicity, suppose that there is a pure action profile a with $g(a) = v$. The proof for the general case follows essentially the same lines. Assume first that the minmax profile m_{-i}^i against each player i is in pure strategies, so that deviations from this profile are certain to be detected. Case ii below sketches how to modify the proof for the case of mixed minmax profiles.

Choose a v' in the interior of V and an $\varepsilon > 0$ such that, for each i,

$$\underline{v}_i < v_i' < v_i,$$

and the vector

$$v'(i) = (v_1' + \varepsilon, \ldots, v_{i-1}' + \varepsilon, v_i', v_{i+1}' + \varepsilon, \ldots, v_I' + \varepsilon)$$

is in V. (The full-dimension assumption ensures that such $v'(i)$ exist for some ε and v'.)

Again, to avoid the details of public randomizations, assume that for each i there is a pure action profile $a(i)$ with $g(a(i)) = v'(i)$. Let $w_i^j = g_i(m^j)$ denote player i's payoff when minmaxing player j. Choose N such that, for all i,

$$\max_a g_i(a) + N\underline{v}_i < \min_a g_i(a) + Nv_i'. \tag{5.8}$$

This is the punishment length such that, for discount factors close to 1, deviating once and then being minmaxed for N periods is worse than getting the lowest payoff once and then N periods of v_i'.

Now consider the following strategy profile:

Play begins in *phase I*. In phase I, play action profile a, where $g(a) = v$. Play remains in phase I so long as in each period either the realized action is a or the realized action differs from a in two or more components. If a single player j deviates from a, then play moves to phase II_j.

Phase II_j Play m^j each period. Continue in phase II_j for N periods so long as in each period either the realized action is m^j or the realized action differs from m^j in two or more components. Switch to phase III_j after N successive periods of phase II_j. If during phase II_j a single player i's action differs from m_i^j, begin phase II_i. (Note that this construction makes sense only if m^j is a pure action profile; otherwise the "realized action" can't be the same as m^j.)

Phase III_j Play $a(j)$, and continue to do so unless in some period a single player i fails to play $a_i(j)$. If a player i does deviate, begin phase II_i.

To show that these strategies are subgame perfect, it suffices to check that in every subgame no player can gain by deviating once and then conforming to the strategies thereafter.

In phase I, player i receives at least v_i from conforming, and he receives at most

$$(1-\delta)\max_a g_i(a) + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i'$$

by deviating once. Since v_i' is less than v_i, the deviation will yield less than v_i for δ sufficiently large. Similarly, if player i conforms in phase III_j, $j \neq i$, then player i receives $v_i' + \varepsilon$. His payoff to deviating is at most

$$(1-\delta)\max_a g_i(a) + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i',$$

which is less than $v_i' + \varepsilon$ when δ is sufficiently large.

In phase III_i, player i receives v_i' from conforming and at most

$$(1-\delta)\max_a g_i(a) + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i'$$

from deviating once. Inequality 5.8 ensures that deviation is unprofitable for δ sufficiently close to 1.

If player i conforms in phase II_j, $j \neq i$, when there are N' periods of phase II_j remaining (including the current period), her payoff is

$$(1-\delta^{N'})w_i^j + \delta^{N'}(v_i' + \varepsilon).$$

If she deviates, she is minmaxed for the next N periods; the play in phase III_j will then give her v_i' instead of the $v_i' + \varepsilon$ she would get in phase III_j if she conformed now. Once again, the ε differential once phase III is reached outweighs any short-term gains when δ is close to 1. Finally, if player i conforms in phase II_i (i.e., when she is being punished) then when there are $N' \leq N$ periods of punishment remaining player i's payoff is

$$q_i(N') \equiv (1-\delta^{N'})\underline{v}_i + \delta^{N'}v_i' < v_i.$$

If she deviates once and then conforms, she receives at most $\underline{v}_i$ in the period in which she deviates (because the opponents are playing m_{-i}^i) and her continuation payoff is then $q_i(N) \leq q_i(N'-1)$.

(ii) The above construction assumes that player i would be detected if she failed to play m_i^j in phase II_j. This need not be the case if m_i^j is a mixed strategy. In order to be induced to use a mixed minmax action, player i must receive the same normalized payoff for each action in the action's support. Since these actions may yield different payoffs in the stage game, inducing player i to mix requires that her continuation payoff be lower after some of the pure actions in the support than after others. Now, in the strategies of part i the exact continuation payoffs for player i in phase III_j, $j \neq i$, were irrelevant (the essential requirement was that player i's payoff be higher in phase III_j than in phase III_i). Thus, as Fudenberg and Maskin (1986a) showed, players can be induced to use mixed actions as punishments by specifying that each player i's continuation payoff in phase III_j, $j \neq i$, vary with the actions player i chose in phase II_j in such a way that each action in the support of m_i^j gives player i the same overall payoff.

As an example of the construction involved, consider a two-player game in which player 1's minmax strategy against player 2 is to randomize $\frac{1}{2}$-$\frac{1}{2}$ between U and D and player 1's payoffs to U and D are 2 and 0, respectively, regardless of how player 2 plays. If player 1 plays U for each of the N periods in phase II_2, he receives an average value of $2(1-\delta^N)$, as opposed to an average value of 0 from playing D and $(1-\delta^N)$ of playing his minmax strategy. So instead of switching to a fixed payoff vector

$$v'(2) = (v_1'(2), v_2'(2))$$

at the end of phase II_2, as in the proof of case i, we specify that player 1's payoff be $v_i'(2) - 2(1-\delta^N)$ if he played U each period, $v_1'(2) - 2\delta(1-\delta^{N-1})$ if he played D at the beginning of phase II_2 and U thereafter, and so on, with the adjustment term chosen so that player 1's average payoff from the start of phase II_2 is $\delta^N v_1'(2)$ for any sequence of phase-II actions that lie in the support of m_1^2. (If player 1 plays an action not in the support of m_1^2, then play switches to phase II_1 as in the proof of i.) ■

Discussion The various folk theorems show that standard equilibrium concepts do very little to pin down play by patient players. In applying repeated games, economists typically focus on one of the efficient equilibria, usually a symmetric one. This is due in part to a general belief that players may coordinate on efficient equilibria, and in part to the belief that cooperation is particularly likely in repeated games. It is a troubling fact that at this point there is no accepted theoretical justification for assuming efficiency in this setting. The concept called "renegotiation proofness," discussed in section 5.4, has been used by a number of authors to reduce the set of perfect-equilibrium outcomes; some versions of this concept imply that behavior must be inefficient.

5.1.3 Characterization of the Equilibrium Set (technical)

The folk theorem describes the behavior of the equilibrium set as $\delta \to 1$. It is also of interest to determine the set of subgame-perfect equilibria for a fixed δ. (The folk theorem suggests that there will be many such equilibria for large discount factors.) Following Abreu (1986, 1988), we will consider the construction of strategies such that any deviation by player i is "punished" by play switching to the perfect equilibrium in which that player's payoff is lowest. In order for this construction to be well defined, we must first verify that these worst equilibria indeed exist.

Theorem 5.5

(i) (Fudenberg and Levine 1983) If the stage game has a finite number of pure actions, there exists a worst subgame-perfect equilibrium $\underline{w}(i)$ for each player i.

(ii) (Abreu 1988) If each player's action space in the stage game is a compact subset of a finite-dimensional Euclidean space, payoffs are continuous for each player i, and there exists a static pure-strategy equilibrium, there is a worst subgame-perfect equilibrium $\underline{w}(i)$ for each player i.

Remarks From the stationarity of the set of equilibria, $\underline{w}(i)$ is also the worst equilibrium in any subgame. The question of whether worst equilibria exist without the pure-strategy restriction in games with a continuum of actions is still open.

Proof

(i) As in chapter 4, with a finite number of actions and payoffs that are continuous at infinity, the set of subgame-perfect equilibria is compact in the product topology on strategies, and payoffs to strategies are continuous in this topology as well. Thus, there are worst (and best) equilibria for each player.

(ii) Let $y(i)$ be the infimum of player i's payoff in any pure-strategy subgame-perfect equilibrium, and let $s^{i,k}$ be a sequence of pure-strategy subgame-perfect equilibria such that $\lim_{k\to\infty} g_i(s^{i,k}) = y(i)$. Let $a^{i,k}$ be the equilibrium path corresponding to strategies $s^{i,k}$, so that

$$a^{i,k} = \{a^{i,k}(0), a^{i,k}(1), \ldots, a^{i,k}(t), \ldots\}.$$

Since A is compact, so is the set of sequences of pure actions (this is Tychonoff's theorem[10]), and we let $a^{i,\infty}$ be an accumulation point. Note that player i's payoff to $a^{i,\infty}$ is $y(i)$.

Now fix a player i, and consider the following strategy profile: Begin in phase I_i.

Phase I_i Play the sequence of actions

$$a^{i,\infty} = \{a^{i,\infty}(0), a^{i,\infty}(1), \ldots\}$$

so long as there are no unilateral deviations from this sequence. If player j unilaterally deviates in period t, then begin phase I_j in period $t+1$. That is, play $a^{j,\infty}(0)$ in period $t+1$, $a^{j,\infty}(1)$ in period $t+2$, and so forth.

If all players follow these strategies, player i's payoff is $y(i)$. To check that the strategies are subgame-perfect, note that if they are not there must exist players i and j, action $\hat{a}_j$, $\varepsilon > 0$, and τ such that

10. See, e.g., Munkres 1975.

$$(1-\delta)g_j(\hat{a}_j, a_{-j}^{i,\infty}(\tau)) + \delta\, y(j) > (1-\delta)\sum_{t=0}^{\infty} \delta^t g_j(a^{i,\infty}(\tau+t)) + 3\varepsilon.$$

Since payoffs are continuous and $a^{i,k} \to a^{i,\infty}$, for k large enough we would have

$$(1-\delta)g_j(\hat{a}_j, a_{-j}^{i,k}(\tau)) + \delta\, y(j) > (1-\delta)\sum_{t=0}^{\infty} \delta^t g_j(a^{i,k}(\tau+t)) + \varepsilon. \tag{5.9}$$

Finally, since $s^{i,k}$ is a subgame-perfect equilibrium, it prescribes *some* subgame-perfect equilibrium if profile $s^{i,k}$ is followed until period τ and then player j plays $\hat{a}_j$ instead of $a_j^{i,k}(\tau)$. Let player j's (normalized) continuation payoff in this equilibrium be $z_j(\tau, \hat{a}_j)$. Since $s^{i,k}$ is subgame perfect,

$$(1-\delta)g_j(\hat{a}_j, a_{-j}^{i,k}(\tau)) + \delta\, z_j(\tau, \hat{a}_j) \le (1-\delta)\sum_{t=0}^{\infty} \delta^t g_j(a^{i,k}(\tau+t)),$$

which contradicts inequality 5.9 since $y(j) \le z_j(\tau, \hat{a}_j)$. ■

Because the players' actions are observed without error, their equilibrium payoffs are not directly affected by the continuation payoffs following actions to which the equilibrium assigns probability 0. Thus, when one is constructing equilibria the magnitudes of such continuation payoffs matter only in that they determine whether or not players can gain by deviating. For this reason, any strategy profile that is "enforced" by some subgame-perfect punishments can be enforced with the harshest punishments available.[11]

Theorem 5.6 (Abreu 1988)

(i) If the stage game is finite, any distribution over infinite histories that can be generated by some subgame-perfect equilibrium σ can be generated with a strategy profile σ^* that specifies that play switches to the worst equilibrium $w(i)$ for player i if player i is the first to play an action to which σ assigns probability 0.

(ii) If the stage game has compact finite-dimensional action spaces and continuous payoffs, then any history $\tilde{h}$ that is generated by a pure-strategy subgame-perfect equilibrium s can be generated by a strategy profile $\hat{s}$ that switches to the worst pure-strategy equilibrium $\underline{w}(i)$ for player i if player i unilaterally deviates from the sequence $\tilde{h}$.

Proof

(i) Fix a perfect equilibrium σ, and construct a new profile σ^* as follows: The profile σ^* agrees with σ (i.e., $\sigma^*(h^t) = \sigma(h^t)$) so long as σ gives the history h^t positive probability. If σ gives positive probability to h^τ for all

11. For readers familiar with the literature on agency, this is the same as the observation that optimal contracts can "shoot the agent" if the observed signal could not have occurred unless the agent cheated.

$\tau < t$, and player i is the only player to play an action outside the support of $\sigma(h^t)$ at period t, then play switches to the the worst subgame-perfect equilibrium for player i, which is $\underline{w}(i)$. More formally,

$$\sigma^*(h^{t-1}) = w(i)(h^0),$$

$$\sigma^*((h^{t-1}, a^{t+1})) \equiv \underline{w}(i)(a^{t+1}),$$

and so on. (As usual, the strategies will ignore simultaneous deviations by two or more players.) Let us verify that σ^* is subgame perfect. In subgames where player i was the first to deviate from the support of σ, σ^* specifies that all players follow profile $w(i)$, which is subgame perfect by definition. In all other subgames h^t the actions prescribed by σ^* are the same as those prescribed by σ, and the continuation payoffs are the same as under σ so long as player i plays an action in the support of $\sigma(h^t)$. It remains to check that player i cannot gain by choosing an action $a_i \notin \text{support}(\sigma_i(h^t))$. If he can, then

$$(1-\delta)g_i(a_i, \sigma_{-i}(h^t)) + \delta u_i(\underline{w}(i)) > u_i(\sigma \mid h^t). \tag{5.10}$$

However, since σ is subgame perfect,

$$u_i(\sigma \mid h^t) \geq (1-\delta)g_i(a_i, \sigma_{-i}(h^t)) + \delta u_i^{\sigma}(a_i \mid h^t), \tag{5.11}$$

where the last term on the right-hand side is player i's continuation payoff from period $t+1$ on under σ if he deviates from σ at h^t by playing a_i. Combining inequalities 5.10 and 5.11 yields the contradiction

$$u_i(w(i)) > u_i^{\sigma}(a_i \mid h^t).$$

(ii) The proof is analogous. ■

Finding the worst possible equilibrium for each player is fairly complicated. However, finding the worst strongly symmetric pure-strategy equilibria of a symmetric game is much simpler, particularly if there are strongly symmetric strategies that generate arbitrarily low payoffs. By "strongly symmetric" we mean that, for all histories h^t and all players i and j,

$$s_i(h^t) = s_j(h^t),$$

so that both players play the same way even after asymmetric histories. For example, in the repeated prisoner's dilemma, the profile where both players use the strategy "tit for tat" (that is, play the action the opponent played the previous period) is not strongly symmetric, since the two players' actions are not identical following the history $h^1 = (\text{C}, \text{D})$. Note that the profile is symmetric in the weaker sense that if $h_1^t = \tilde{h}_2^t$ and $h_2^t = \tilde{h}_1^t$, then

$$s_1(h_1^t, h_2^t) = s_2(\tilde{h}_1^t, \tilde{h}_2^t),$$

so that permuting the past history permutes the current actions. We use the terms "strongly symmetric" and "symmetric" to distinguish between these two notions of symmetry.

Abreu (1986) shows that the worst strongly symmetric equilibrium is very easy to characterize in symmetric games where the action spaces are intervals of real numbers, payoffs are continuous and bounded above, (a) the payoff to symmetric pure-strategy profiles $\vec{a}$ (i.e., profiles where each player i plays a) is quasiconcave, and decreases without bound as a tends to infinity, and (b) letting $\vec{a}_{-i}$ denote the profile in which all of player i's opponents choose action a, the maximal payoff to deviating from pure-strategy symmetric profile $\vec{a}$,

$$\max_{a_i'} g_i(a_i', \vec{a}_{-i}),$$

is weakly decreasing in a.

Condition b is natural in a symmetric quantity-setting game, where by producing very large outputs firms drive the price to 0 and thus make the best payoff to deviating very small.

We emphasize that in the definition of a strongly symmetric equilibrium, symmetry is required off the equilibrium path as well as on the path, which rules out many asymmetric punishments that could be used to enforce symmetric equilibrium outcomes.

Theorem 5.7 (Abreu 1986) Consider a symmetric game satisfying conditions a and b. Let e^* and e_* denote the highest and lowest payoff per player in a pure-strategy strongly symmetric equilibrium.

(i) The payoff e_* can be attained in an equilibrium with strongly symmetric strategies of the following form: "Begin in phase A, where the players play an action a_* that satisfies

$$(1 - \delta)g(\vec{a}_*) + \delta e^* = e_*. \tag{5.12}$$

If there are any deviations, continue in phase A. Otherwise, switch to a perfect equilibrium with payoffs e^* (phase B)."

(ii) The payoff e^* can be attained with strategies that play a constant action a^* as long as there are no deviations, and switch to the worst strongly symmetric equilibrium if there are any deviations. (Other feasible payoffs can be attained in a similar way.)

Proof

(i) Fix some strongly symmetric equilibrium $\hat{s}$ with payoff e_* and first-period action a. Since the continuation payoffs under $\hat{s}$ cannot be more than e^*, the first-period payoffs $g(\vec{a})$ must be at least $(-\delta e^* + e_*)/(1 - \delta)$. Thus, under condition a there is an $a_* \geq a$ with $g(\vec{a}_*) = (-\delta e^* + e_*)/(1 - \delta)$. Let s_* denote the strategies constructed in the statement of the theorem. By definition, the strategies s_* are subgame perfect in phase B. In phase A, condition b and $a_* \geq a$ imply that the short-run gain to deviating is no more than that in the first period of $\hat{s}$. Since the punishment for deviating in phase A is the worst possible punishment, the fact that no player preferred to deviate in the first period of $\hat{s}$ implies that no player prefers to deviate in phase A of s_*.

(ii) We leave the proof of part ii to the reader. ■

Remark When this theorem applies, the problem of characterizing the best strongly symmetric equilibrium reduces to finding two numbers, representing the actions in the two phases. (One application is given by Lambson (1987).) If the action space is bounded above by some $\bar{a}$, payoffs cannot be made arbitrarily low, and the punishment phase A may have to last for several periods. In this case it is not obvious precisely which actions should be specified in phase A. The obvious extension of the theorem would have the players using $\bar{a}$ for T periods and then, as before, switching to phase B, where the continuation payoff is e^*. The difficulty is that there may not be a T such that the resulting payoffs in phase A, which are $(1 - \delta^T)g(\bar{a}) + \delta^T e^*$, exactly equal e_*, as required by equation 5.12. However, if we assume that public randomizing devices are available, this integer problem can be eliminated. (Remember that for small discount factors the public-randomization assumption can change the set of equilibria.)

Abreu also shows that in general the highest symmetric pure-strategy equilibrium payoff requires "punishments" with payoffs e_* less than the payoffs in any static equilibrium, unless the threat of switching to the static equilibrium forever—i.e., the strategies introduced by Friedman—supports an efficient outcome.

Finally, Abreu shows that under conditions a and b symmetric pure-strategy equilibria support payoffs on the frontier of the equilibrium set if and only if there is a strongly symmetric equilibrium that gives players their minmax values.

Fudenberg and Maskin (1990b) consider stage games with finitely many actions. They observe that when for each player i there is a perfect equilibrium in which player i's payoff is $\underline{v}_i$, the sets of Nash-equilibrium and perfect-equilibrium payoffs coincide and provide conditions in the stage game for which such perfect equilibria exist for all sufficiently large discount factors. (See exercise 5.8.)

5.2 Finitely Repeated Games^{+++}

These games represent the case of a fixed known horizon T. The strategy spaces at each $t = 0, 1, \ldots, T$ are as defined above; the utilities are usually taken to be the time average of the per-period payoffs. (Allowing for a discount factor δ close to 1 will not change the conclusions we present.)

The set of equilibria of a finitely repeated game can be very different from that of the corresponding infinitely repeated game, because the scheme of self-reinforcing rewards and punishments used in the folk theorem can unravel backward from the terminal date. The classic example of this is the repeated prisoner's dilemma. As observed in chapter 4, with a fixed finite horizon "always defect" is the only subgame-perfect-equilibrium outcome. In fact, with a bit more work one can show this is the only *Nash* outcome:

Fix a Nash equilibrium σ^*. Both players must cheat in the last period, T, for any history h^T that has positive probability under σ^*, since doing so increases their period-T payoff and since there are no future periods in which they might be punished. Next, we claim that both players must defect in period $T-1$ for any history h^{T-1} with positive probability: We have already established that both players will cheat in the last period along the equilibrium path, so in particular if player i conforms to the equilibrium strategy in period $T-1$ his opponent will defect in the last period, and hence player i has no incentive not to defect in period $T-1$. An induction argument completes the proof. This conclusion, though not pathological, relies on the fact that the static equilibrium gives the players exactly their minmax values, as the following theorem shows.

Theorem 5.8 (Benoit and Krishna 1987) Assume that for each player i there is a static equilibrium $\alpha^*(i)$ with $g_i(\alpha^*(i)) > \underline{v}_i$. Then the set of Nash-equilibrium payoffs of the T-period game with time averaging converges to the set of feasible, individually rational payoffs as $T \to \infty$.

Proof The key idea of the proof is to first construct a "terminal reward phase" in which each player receives strictly more than his minmax value for many periods. To do this, let the "reward cycle" be the sequence of mixed-action profiles $\alpha^*(1), \alpha^*(2), \ldots, \alpha^*(I)$, and let the R-cycle terminal phase be the sequence of profiles of length $R \cdot I$ where the reward cycle is repeated R times. Any R-cycle terminal phase is clearly a Nash-equilibrium path in any subgame of length $R \cdot I$. And since each $\alpha^*(j)$ gives player i at least his minmax value and $\alpha^*(i)$ by assumption gives him strictly more, each player's average payoff in this phase strictly exceeds his minmax level.

Next, fix a feasible, strictly individually rational payoff v, and set R large enough so that each player i prefers payoff v_i followed by the R-cycle terminal phase to getting the largest possible payoff, $\max_a g_i(a)$, in one period and then being minmaxed for $R \cdot I$ periods. Then choose any $\varepsilon > 0$, and choose T so that there is a deterministic cycle of pure actions $\{a(t)\}$ of length $T - R \cdot I$ whose average payoffs are within ε of payoff v.

Finally, we specify the following strategies: Play according to the deterministic cycle $\{a(t)\}$ in each period so long as past play accords with $\{a(t)\}$ and there are more than $R \cdot I$ periods left. If any player unilaterally deviates from this path when there are more than $R \cdot I$ periods left, then minmax that player for the remainder of the game. If play agrees with $\{a(t)\}$ until there are $R \cdot I$ periods left, then follow the R-cycle terminal phase for the remainder of the game regardless of the observed actions in this phase.

These strategies are a Nash equilibrium for any $T > R \cdot I$. For $T > R \cdot I(\max_a g_i(a) - v_i)/\varepsilon$, the average payoffs are within 2ε of v. ■

Benoit and Krishna (1985) give a related result for subgame-perfect equilibria under a stronger condition. (Friedman (1985) and Fraysse and Moreaux (1985) give independent, less complete analyses of special classes of games.) Recall from chapter 4 that if the stage game has a unique equilibrium, backward induction shows that the unique perfect equilibrium of the finitely repeated game is to play the static equilibrium in every period of every subgame. Where there are several static equilibria, it is possible to punish a player for deviating in the next-to-last period by specifying that if he does not deviate the static equilibrium he prefers will occur in the last period, and that deviations lead to the static equilibrium he likes less.

Theorem 5.9 (Benoit and Krishna 1985) Assume that for each player i there are static equilibria $\alpha^*(i)$ and $\hat{\alpha}(i)$ with $g_i(\alpha^*(i)) > g_i(\hat{\alpha}(i))$, and that the dimension of the feasible set equals the number of players. Then, for every feasible payoff $v \in V$ with v_i strictly exceeding player i's pure-strategy minmax level, and for every sufficiently small $\varepsilon > 0$, there is a T such that for all finite horizons $T' > T$ there is a subgame-perfect equilibrium whose payoffs are within ε of v.

Proof Omitted.

As in the infinite-horizon case, the full-dimension condition is needed to allow strategies that reward one player without rewarding another. The question of whether this result can be strengthened to obtain all payoffs above the mixed-strategy minmax levels is still open.

Although the Benoit-Krishna results extend the Nash-equilibrium and perfect-equilibrium folk theorems to a class of finitely repeated games, in games like the prisoner's dilemma the only Nash equilibrium with finite repetitions is to always be "unfriendly." Few "real-world" long-term relationships correspond to the finite-horizon model; however, there have been many experimental studies of games in which the participants are indeed told that the horizon has been set at a fixed finite point, and there is a unique stage-game equilibrium. In such experimental studies of the prisoner's dilemma, subjects do in fact tend to cooperate in many periods, despite what the theory predicts.

One response is that players are known to derive some extra satisfaction from "cooperating" above and beyond the rewards specified in the experimental design. This explanation does not seem implausible, but it is a bit too convenient, and seemingly much too powerful; once we admit the possibility that payoffs are known to be misspecified, it is hard to see how any restrictions on the predicted outcome of the experiment could be obtained.

A second response is to make a smaller change in the model and allow for there to be a *small probability* that the players get extra satisfaction from cooperating, *so long as* their opponent has cooperated with them in the past. This is the basis of the Kreps-Milgrom-Roberts-Wilson (1982) idea of "reputation effects," which we discuss in detail in chapter 9. The ε-equilibrium approach (Radner 1980; Fudenberg and Levine 1983), discussed in section 4.8, is another way of derailing backward induction in the finitely repeated game, although this requires adopting ε-equilibrium as a descriptive model of bounded rationality rather than merely a convenient technical device.

5.3 Repeated Games with Varying Opponents†††

Classic repeated games suppose that the same fixed set of players play one another every period. However, results similar to the folk theorem can be obtained in some cases where not all of the players play one another infinitely often. This section discusses several variants of this idea.

5.3.1 Repeated Games with Long-Run and Short-Run Players

The first variant we will consider supposes that some of the players are long-run players, as in standard repeated games, while the roles corresponding to other "players" are filled by a sequence of short-run players, each of whom plays only once.

Example 5.1

Suppose that a single long-run firm faces a sequence of short-run consumers, each of whom plays only once but is informed of all previous play when choosing his actions. Each period, the consumer moves first, and chooses whether or not to purchase a good from the firm. If the consumer does not purchase, then both players receive a payoff of 0. If the consumer decides to purchase, then the firm must decide whether to produce high or low quality. If it produces high quality, both players have a payoff of 1; if it produces low quality, the firm's payoff is 2 and the consumer's payoff is -1. This game is a simplified version of those considered by Dybvig and Spatt (1980), Klein and Leffler (1981), and Shapiro (1982).[12] Simon (1951)

12. Dybvig and Spatt (1980) and Shapiro (1982) consider models where the place of the "short-run player" described above is taken by a continuum of long-lived but "small" consumers. Since they assume that the play of any individual consumer cannot be observed, the consumers always act to maximize their current payoff, and the models are equivalent to the case of a sequence of individual, short-run consumers. (See our treatment of open- and closed-loop equilibrium in chapter 3 for a discussion of the assumption that the play of "small" players cannot be observed.)

and Kreps (1986) use a similar game to analyze the employment relationship, and to argue that one reason for the existence of "firms" is precisely to provide a long-run player who can be induced to be trustworthy by the prospect of future rewards and punishments.

The following strategies are a subgame-perfect equilibrium of this game when the firm is sufficiently patient: The firm starts out producing high quality every time a consumer purchases, and continues to do so as long as it has never produced low quality in the past. If ever the firm produces low quality, it produces low quality at every subsequent opportunity. The consumers start out purchasing the good from the firm, and continue to do so so long as the firm has never produced low quality. If ever the firm produces low quality, then no consumer ever purchases again. The consumer's strategies are optimal because each consumer cares only about that period's payoff, and thus should buy if and only if that period's quality is expected to be high. The firm does incur a short-run cost by producing high quality, but when the firm is patient this cost is offset by the fear that producing low quality will drive away future consumers. Note that this equilibrium suggests why consumers may prefer to deal with a firm that is expected to remain in business for a while, as opposed to a "fly-by-night" firm for whom long-run considerations are unimportant.

Example 5.2

As a second example, consider repeated play of a sequential-move version of the prisoner's dilemma with a single long-run player facing a sequence of short-run opponents. Each period, the short-run player's decision whether to cooperate or to cheat is observed before the long-run player makes his own decision. As in the previous example, if the long-run player's discount factor is close to 1, there is an equilibrium where players always cooperate. One such equilibrium is: "The short-run players cooperate so long as in every past period the long-run player has played the same action as that period's short-run player; if the long-run player has ever failed to match the short-run players cheat. The long-run player matches the play of that period's opponent so long as he has never failed to match in the past, and cheats otherwise."

The key to the cooperative equilibrium in example 5.2 is that since the short-run players move first, they can be provided with an incentive to cooperate without the use of rewards and punishments in future periods. If instead the moves in the stage game are simultaneous, the short-run players will cheat in every period, and so the only equilibrium outcome is for both sides to always cheat. This suggests that the way to extend the folk theorem to these games is to modify the definitions of the feasible payoffs and the minmax values to incorporate the constraint that short-run players always play short-run best responses.

To state this conjecture formally, label the players so that players $i = 1, \ldots, \ell$ are long-run players who maximize the normalized discounted sum of their per-period payoffs as in ordinary repeated games, and let players $j = (\ell + 1), \ldots, I$ represent sequences of short-run players who act in each period to maximize that period's payoff. That is, the stage game has I players, and in the repeated game the individuals playing the parts of players $\ell + 1$ through I change each period. (Alternatively, players $\ell + 1$ to I could be long-run players whose discount factor is 0.) Let

$$\mathrm{B}\colon \mathcal{A}_1 \times \cdots \times \mathcal{A}_\ell \to \mathcal{A}_{\ell+1} \times \cdots \times \cdots \times \mathcal{A}_I$$

be the correspondence that maps any action profile $(\alpha_1, \ldots, \alpha_\ell)$ for the long-run players to the corresponding Nash-equilibrium actions for the short-run players. That is, for each $\alpha \in \text{graph}(\mathrm{B})$ and $i \geq \ell + 1$, α_i is a best response to α_{-i}.

For each long-run player i, define the minmax value $\underline{v}_i$ to be

$$\min_{\alpha \in \text{graph}(\mathrm{B})} \; \max_{a_i \in A_i} g_i(a_i, \alpha_{-i}). \tag{5.13}$$

(The minimum is attained because the graph of B is compact, and the payoff functions are continuous in the mixed strategies. Note that this definition reduces to the usual one if all the players are long-run.) Let

$$U = \{v = (v_1, \ldots, v_\ell) \in \mathbb{R}^\ell \mid \exists \alpha \in \text{graph}(\mathrm{B}) \text{ with } g_i(\alpha) = v_i \text{ for } i = 1, \ldots, \ell\}$$

and set

$$V = \text{convex hull}\,(U).$$

This is the modified definition of the set of feasible payoffs.

As we remarked, one might suspect that the folk theorem would extend with these modified definitions of feasibility and the minmax levels. However, as shown by Fudenberg, Kreps, and Maskin (1990) this extension obtains only when each player's choice of a mixed action in the stage game is publicly observable. When players observe only their opponents' realized actions, the set of subgame-perfect equilibria can be strictly smaller. The reason for this, as illustrated in exercise 5.9, is that, in order to induce a short-run player to take a particular action along the equilibrium path, some of the long-run players may need to use mixed actions. When the randomizing probabilities are not observable, inducing this randomization will require that the continuation payoffs make the randomizing long-run players indifferent between the pure actions they assign positive probability, which imposes a cost in terms of the efficiency of the possible equilibrium payoffs.

The limit set of equilibria with unobserved randomizing probabilities is the intersection of the feasible, individually rational payoffs with the constraints $v_i \leq \bar{v}_i$, where $\bar{v}_i$ is defined as

$$v_i = \max_{\alpha \in \text{graph}(B)} \min_{a_i \in \text{support}(\alpha_i)} g_i(a_i, \alpha_{-i}). \tag{5.14}$$

For a fixed mixed-action profile α, equation 5.14 computes player i's worst payoff among the actions α_i requires him to play with positive probability. Intuitively, if player i is asked to play α_i along the equilibrium path, he must be willing to use every action in α_i.

Theorem 5.10 (Fudenberg, Kreps, and Maskin 1990; Fudenberg and Levine 1990) Assume that the dimension of V is equal to ℓ, the number of long-run players. Then for every $v \in V$ with $\underline{v}_i < v_i < \overline{v}_i$ for all $i = 1, \ldots, \ell$, there is a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a subgame-perfect equilibrium with payoffs v.

Proof Omitted.

5.3.2 Games with Overlapping Generations of Players

Crémer (1986) considered a repeated game in which overlapping generations of players live for T periods, so that at each date t there is one player of age T who is playing his last round, one player of age $T - 1$ who has two rounds still to play, and so on down to the new player who will play T times. Each period, the T players simultaneously choose whether to work or to shirk, and their choices are revealed at the end of each period; players share equally in the resulting output, which is an increasing function of the number who chose to work.[13] The cost of effort exceeds a $1/T$ share of the increases in output, so shirking is a dominant strategy in the stage game, which has the flavor of a T-player prisoner's dilemma. Payoffs in the repeated game are the average of the per-period utilities.

Suppose that the efficient outcome is for all players to work. This outcome cannot occur in any Nash equilibrium, since the age-T player will always shirk. Nevertheless, there can be equilibria where most of the players work. This will be easiest to see if we further specialize the model. Let $T = 10$. Suppose that if k players work the aggregate output is $2k$, and that the disutility of effort is 1. Then if preferences are linear in output and effort, the payoff to working when k opponents work is $2(k + 1)/10 - 1$, and the payoff to shirking is $2k/10$. The efficient outcome is for all players to work, with resulting utility of 1 per player.

Now consider the following strategy profile: "Age-10 players always shirk. So long as no player has ever shirked when his age is less than 10, all players of age less than 10 work. If a player has ever shirked when his age is less than 10, then all players shirk." If all players conform to this profile, each player receives $18/10 - 1 = 4/5$ in the periods he works and

13. It is, however, interesting to note that the "cooperative" equilibrium we derive in the next paragraph remains an equilibrium if we suppose that workers observe only the total number of shirkers but not their identities.

9/5 in the period he is of age 10. Clearly, no player can gain by deviating when he is of age 10. If a player of age 9 deviates, he receives 8/5 the period he deviates, and 0 the next period, which is less than 4/5 + 9/5; younger players lose even more by deviating. Thus, these strategies are a subgame-perfect equilibrium.

Kandori (1989b) and Smith (1989) have generalized this type of construction and provided conditions for the folk theorem to obtain.

5.3.3 Randomly Matched Opponents

Another variant of the repeated-games model supposes that there are *a* many players, each of whom plays infinitely often but against a different opponent each period. More precisely, fix a two-player stage game, and suppose that there are two populations of players of equal size, N. Each period, every player 1 is matched with a player 2. The probability of being matched to a particular player 2 is $1/N$, and matching in each stage is independent.[14]

In the first analyses of this sort of random-matching model, Rosenthal (1979) and Rosenthal and Landau (1979) assumed that when the players in each pair are matched, their information consists of the actions that the two of them played in the previous period. Thus, if the stage game is the prisoner's dilemma, where C is "cooperate" and D is "defect," there are four possible "histories" a pair of players can have, namely (C, C), (D, C), (C, D), and (D, D), and consequently each player has $2^4 = 16$ pure strategies. (Note that players do not have perfect recall!)

With this information structure, the strategy "cooperate if and only if my opponent cooperated last period," or "tit for tat," is feasible. More generally, the action a player chooses in period t can have a direct effect on his opponent's play in period $t + 1$.

If the player expects to face the same opponent in period $t + 1$ and in period $t + 2$, he may anticipate an additional indirect effect of his period-t action on his opponent's play in periods after $t + 1$. For example, if the opponent's strategy is to cooperate only if the history is (C,C), defecting in period t will not only make the opponent defect in period $t + 1$; it will also make the opponent defect in every period thereafter.

Rosenthal (1979) and Rosenthal and Landau (1979) restrict their attention to "Markovian equilibria," where this indirect effect is not present and where each player believes that his action at date t has no effect on the play of his opponent at all dates from $t + 2$ on.[15] Although this belief is incorrect with a single player of each type, it is correct in a model with a continuum

14. This kind of model can be used to explain why, e.g., traders may behave honestly even though it is very unlikely that they will ever meet each other again in the future (Greif 1989; Milgrom, North, and Weingast 1989).

15. The Markovian notion here differs from the one defined in chapter 13.

of each kind of player, so that no player ever meets the same opponent twice.[16]

When is all players using "tit for tat" a Markovian equilibrium of the prisoner's dilemma? Each player must be willing to cooperate if the current opponent cooperated last period, and must be willing to defect if the opponent defected last period. Yet, the next period's opponent will not know the past play of the current opponent, and thus cannot distinguish between a "defect" that occurred to punish the current opponent for a past defection and a "defect" that represents a deviation from the strategy "tit for tat." In particular, with the strategy "tit for tat," any defection today will make the next opponent defect. Thus, both players using "tit for tat" is a Markovian equilibrium only if the discount factor is exactly such that the short-run gain to cheating exactly equals the discounted cost of being punished next period: If the discount factor is smaller, then the threat of punishment will not enforce cooperation; if the discount factor is larger, then a player whose opponent defected last period will not be willing to punish him, as doing so will reduce the punisher's future payoff. With payoff as in figure 4.1, this critical value is $\delta = \frac{1}{2}$. More generally, Rosenthal shows that for all but one value of the discount factor, the unique symmetric Markovian equilibrium of the prisoner's dilemma is for all players to cheat in every period. (Exercise 5.7 asks you to check this.)

Kandori (1989b) observes that cooperation *is* an equilibrium outcome for discount factors near 1 if each player observes the *outcome* in his partner's previous match, i.e., the play of both his partner and the partner's opponent. In this case cooperation can be enforced by the strategies "Cooperate in the first period, and continue to cooperate as long as the outcome in each of my matches has been (C, C), and the outcome in my opponent's last match was (C, C); otherwise defect." With these strategies a player whose partner cheated last period can do no better than carry out the prescribed punishment, as the current partner will defect this period, and so the player will be punished next period regardless of how he plays today. Also, a player who deviates will be punished forever, regardless of his future play. Kandori notes that these strategies have the unappealing feature that a deviation by a single player causes the whole "society" to eventually unravel to the all-defect equilibrium. He proposes that researchers should look for equilibria that are "resilient" in the sense that play will eventually return to cooperation in any subgame (that is, after any finite sequence of deviations). Since the motivation for this kind of stability is the idea that there may be some noise in the model that triggers the "punishment

16. To avoid technical complications, Rosenthal (1979) and Rosenthal and Landau (1979) suppose that each population of players is finite, so that there is a nonzero chance that a player 1 will be matched with the same player 2 in two successive periods. Thus, player 1's action today may in fact have some influence on the way his next opponent will play, but with many players in each population this influence is small.

scheme," an alternative methodology would be to make the noise explicit. This would transform the prisoner's dilemma into a game with imperfectly observed actions, a topic we discuss in sections 5.5–5.7. Studying random-matching equilibria in games with noise is an open problem in the literature at this time.

Kandori also suggests another type of equilibrium for games in which players observe only the play in their own past contests. In this "contagion" equilibrium, all players initially cooperate, and if a player ever encounters an opponent who plays D, he plays D from then on. With an infinite population of players, so that with probability 1 no player will ever meet his current opponent again (nor will he even meet anyone who has played his current opponent, etc.), players have no long-run loss from playing D, and these strategies are not an equilibrium. However, with a finite population and random matching, playing D today will eventually lead the entire population to play D. Thus, there is the potential for the contagion strategies to be an equilibrium; whether they are or not depends on how fast the contagion spreads, which in turn depends on the number of players. If there are only two players, the contagion strategies are clearly an equilibrium for discount factors close to 1. For fixed stage-game payoffs, Kandori's contagion strategies fail to be an equilibrium, but not because players are tempted to defect in the cooperative phase. Rather, the problem is that players prefer to continue playing C even after meeting an opponent who plays D in order to slow down the spread of the contagion process. Kandori shows that, for any fixed number of players, the contagion strategies are an equilibrium for discount factors close to 1 provided that the payoffs in the stage game are altered to make the payoff to playing C against an opponent who plays D sufficiently negative. In this case, even a very small probability that the next opponent plays D is sufficient to make D the best response.

Ellison (1991) shows that for any number of players and fixed stage-game payoffs there is in fact an equilibrium where all players cooperate. Moreover, this equilibrium is partially resilient, in the sense that if a single player cheats once the resulting steady state is for players to continue to cooperate most of the time. (The equilibrium can be made entirely resilient if public randomizing devices are introduced.) Ellison also constructs an approximately efficient equilibrium of the random-matching model with noise.

5.4 Pareto Perfection and Renegotiation-Proofness in Repeated Games†††

5.4.1 Introduction

Recently many economists have studied the idea of the "renegotiation" of equilibria and, in particular, the consequences of such renegotiation for play in repeated games. The idea is that if equilibrium arises as the result of negotiations between the players, and players have the opportunity to negotiate anew at the beginning of each period, then equilibria that enforce "good" outcomes by the threat that deviations will trigger a "punishment equilibrium" may be suspect, as a player might deviate and then propose abandoning the punishment equilibrium for another equilibrium in which all players are better off. This sort of equilibrium restriction is called "Pareto perfection" because it extends the idea that players will not play a Pareto-dominated equilibrium to dynamic settings by requiring that in any subgame the equilibrium played must not be Pareto dominated given the constraints on the equilibria at future dates.

The restriction is also called "renegotiation-proofness," because the constraint of Pareto optimality in the subgames can be interpreted as the result of the players' "renegotiating" the original agreement. This latter terminology suggests a parallel with the literature on the renegotiation of contracts, which has also developed a notion of "renegotiation-proofness," but the parallel is inexact: If two players agree on a contract, its terms are legally binding unless both players agree to replace the contract with another one; in contrast, the original "negotiations" on an equilibrium are not binding and serve only to coordinate expectations.

Since the process of selecting a Pareto-optimal outcome and the ideas of Pareto perfection and renegotiation-proofness all take as their starting point the premise that in a static game players will always play an equilibrium on the Pareto frontier of the set of equilibrium payoffs, they are subject to the various critiques of that assumption. In particular, consider the game illustrated in figure 5.3, which we discussed in subsection 1.2.4.

We argued in subsection 1.2.4 that even though equilibrium (U, L) Pareto dominates the others, it is not clear that it is the most reasonable prediction of how the game will be played, even if the players can communicate before the game is played. As Aumann (1990) observes, regardless of his own play, player 2 gains if player 1 plays U, and so regardless of how player 2 intends to play he should tell player 1 that he intends to play L. Thus, it is not clear that the players should expect that their opponent believes that their announcements are sincere.

The concepts developed in this section go further and suppose that even if players have deviated from the play prescribed by past "negotiations," future negotiations will have an efficient outcome. For example, in a twice-

	L	R
U	9,9	0,8
D	8,0	7,7

Figure 5.3

repeated version of the game illustrated in figure 5.3, Pareto perfection requires that players play (U, L) in the first period, and that they play (U, L) in the second period even if one or both of them deviated in the first period — the first-period deviation is treated as a "bygone" that has no effect on subsequent play. This is a strong assumption. It can, however, be rationalized if we suppose that, as in subgame perfection, players treat deviations as accidents that are unlikely to be repeated. (Chapter 11 discusses the idea of forward induction, where deviations are interpreted as strategic signals; this yields quite different conclusions about repeated games.)

Despite our reservations about Pareto efficiency in static games, the concept is sufficiently interesting that we want to discuss its dynamic counterpart. There are currently several competing theories of what this dynamic counterpart should be; the folk theorem obtains under some of them but not others, as we explain below. We begin with the case of finitely repeated games, where it is easier to see what the "right" definitions should be.

5.4.2 Pareto Perfection in Finitely Repeated Games

The best-established formal definition of a renegotiation-proof equilibrium concept is the Bernheim-Peleg-Whinston (1987) notion of a Pareto-perfect equilibrium of a finitely repeated game. Pareto perfection combines the ideas of the Pareto optimality of equilibrium with the logic of subgame perfection, resulting in the recursive definition given below.[17]

For any set C in $\mathbb{R}^I$, let Eff(C) be the set of strongly efficient points in C, i.e., the set of $x \in C$ such that there is no $y \in C$ with $y \geq x$ and $y \neq x$.

Definition 5.1 (Bernheim, Peleg, and Whinston 1987) Fix a stage game g, and let G^T be the associated T-fold repetition. Let P^T be the set of payoffs of pure-strategy subgame-perfect equilibria of G^T. Set $Q^1 = P^1$ and $R^1 = \text{Eff}(P^1)$.

For $T > 1$, let $Q^T \subseteq P^T$ be the set of pure-strategy perfect-equilibrium payoffs that can be enforced with continuation payoffs in R^{T-1} in the second period of the game, and set $R^T = \text{Eff}(Q^T)$.

A perfect equilibrium σ of G^T is Pareto perfect if, for every time t and history h^t, the continuation payoffs under σ are in R^{T-t}.

The restriction to pure-strategy equilibria is commonly used in this literature. However, since some games have mixed-strategy equilibria that Pareto dominate all the pure-strategy equilibria, this restriction is not innocuous. Also, recall from subsection 1.2.4 that "negotiation"-type argu-

17. Bernheim, Peleg, and Whinston (1987) give a definition for general multi-stage games with observed actions. We specialize to repeated games for notational convenience but the general definition should be clear. Our formulation of the definition is taken from Benoit and Krishna (1988), who restrict their attention to pure-strategy equilibria.

	b_1	b_2	b_3	b_4
a_1	0,0	2,4	0,0	5.5,0
a_2	4,2	0,0	0,0	0,0
a_3	0,0	0,0	3,3	0,0
a_4	0,5.5	0,0	0,0	5,5

Figure 5.4

ments support Pareto-optimal equilibria only in two-player games. Though Bernheim, Peleg, and Whinston extend their concept of coalition-proof equilibrium to that of perfectly coalition-proof equilibrium, most subsequent work has focused on two-player games.

To see the force of renegotiation constraints, consider the example illustrated in figure 5.4, which was used by Benoit and Krishna (1988) and by Bergin and MacLeod (1989). In this game the set R^1 of Pareto-optimal pure-strategy equilibrium payoffs is $\{(4,2),(3,3),(2,4)\}$. Because there are multiple elements in R^1, in the twice-repeated game G^2 we are free to vary the last-period equilibrium with the first-period play. This allows the payoffs $(5,5)$ to be enforced in the first period if the players are sufficiently patient: Specify that the continuation will be $(3,3)$ if there are no deviations, and that a deviator will be punished with a continuation payoff of 2. In particular, if (as Benoit and Krishna assume) the discount factor is exactly 1, then R^2 is the single point $(8,8)$. But now in G^3 there is no way to enforce cooperation in the first period, as the continuation game G^2 has a unique Pareto-perfect payoff! Thus, Pareto perfection requires that one of the static equilibria occur in the first period of the thrice-repeated game, and $Q^3 = R^3 = \{(12,10),(11,11),(10,12)\}$. Given the variation in the payoffs allowed by Q^3, profile (a_4, b_4) can be enforced in the first period of a four-stage game, and so on. Benoit and Krishna show that the sets R^T alternate, having three elements if T is odd and a single element if T is even. Moreover, as $T \to \infty$, the average payoffs per period, R^T/T, converge to the point $(4,4)$, even though from Benoit and Krishna 1985 the efficient payoffs $(5,5)$ can be approximated in a subgame-perfect equilibrium when T is large (see section 5.2). Thus, the Pareto-perfect equilibria need not be Pareto efficient in the set of all perfect equilibria, as the restriction to Pareto-perfect continuations reduces the frequency with which players can be induced to play the efficient pair (a_4, b_4).

Note an interesting way in which the set R^T differs from the set P^T of all perfect equilibria: Even with a very long horizon, the play in the first few periods is very sensitive to the exact period length. Thus, the assumption of a precisely known horizon is even more important here than when the renegotiation constraint is not imposed, for then (under the conditions of

Benoit and Krishna (1985)), with a long horizon, play until the "last few periods" need not depend on the exact length of the game.

Benoit and Krishna prove that for general stage games the set of average payoffs R^T/T converges either to a single point, as in the example, or to a subset of the efficient frontier. More precisely, they prove that these properties hold when the recursive definition of R^T is modified to consider at each stage T only the pure-strategy equilibria with continuations in R^{T-1}. (Recall that there is a sense in which this restriction to pure-strategy equilibria conflicts with the criterion of Pareto optimality, as there are stage games in which all pure-strategy equilibria are Pareto dominated by equilibria in mixed strategies.)

Bergin and MacLeod (1989) offer an alternative to renegotiation-proofness for finitely repeated games that they call *recursive efficiency*. Recursive efficiency is defined recursively like Pareto perfection, with Q'^T as the equilibrium payoffs enforceable with continuations in R'^{T-1}; the difference is that the set R'^T of recursively efficient agreements is allowed to be a proper subset of the efficient points of Q'^T.[18]

In the example above, recursive efficiency allows the set R'^1 to be the singleton (3, 3), so that (5, 5) could not be enforced in the first period of G^2. This, in turn, allows R'^2 to be the set $\{(5,7),(6,6),(7,5)\}$, so that in G^3 the outcome (a_4, b_4) with payoffs (5, 5) can be enforced in period 1 by specifying that if there are no deviations the continuation payoffs are (6, 6) while a deviator receives continuation payoff of 5. This contrasts with Pareto perfection, which requires that the continuation payoffs from period 2 on be (8, 8), and thus precludes "cooperation" in the first period.

If the discount factors are exactly 1, this time shift does not affect the players' payoffs in G^3, but if the discount factor is less than 1 the players prefer to have the high payoffs (5, 5) occur in the first period. For example, if $\delta = \frac{1}{2}$ the discounted Pareto-perfect payoffs in G^3 are (25/4, 25/4) and the recursively efficient payoffs are (29/4, 29/4). (If the discount factor is too small, the strategies described are not perfect.)

Bergin and MacLeod justify their alternative definition as follows. Suppose that the players meet before period 1 and agree to play any fixed static equilibrium in the last period, and that this continuation equilibrium then becomes the "social norm." Moreover, all players believe that, regardless of the negotiations in period 2, the social norm for period 3 will be played unless it is unacceptable *once period 3 is reached*. Then in period 2, the suggestion to play the more efficient equilibrium "(a_4, b_4) today, any deviator gets 2 tomorrow" is not credible, even though it is Pareto perfect, as both players would feel free to deviate and then appeal to the "social

18. Recursive efficiency also replaces efficiency with weak efficiency. (For any set C in $\mathbb{R}^I$, Weff(C), the set of weakly efficient points in C is the set of $x \in C$ such that there is no $y \in C$, with $y > x$; i.e., y dominates x in all its components.)

norm" of (3, 3) in the last period. In other words, recursive efficiency gives some weight to the original agreement, while under Pareto perfection the set of original agreements is not considered when choosing agreements for the continuation game. As Bergin and MacLeod put it, Pareto perfection is "a theory for which history is not important," while under recursive efficiency "the agreement in period 1 acts as a default focal point." From another viewpoint, recursive efficiency supposes "less superrationality" at the renegotiation stage, as players cannot renegotiate to the Pareto-perfect equilibrium of the last two periods.

The spirit of the Bergin-MacLeod interpretation can be used to justify a less restrictive notion of recursive efficiency. Bergin and MacLeod allow the set of agreements at time t to be a subset of the efficient agreements relative to the recursively efficient continuations, but do not allow the subset Q^t chosen to depend on the prior history h^t. Consider a four-period version of the game illustrated in figure 5.4, with discount factor $\frac{1}{2}$. If the players can agree to use an equilibrium in periods 2 and 3 that plays (a_1, b_2) even though it is not Pareto perfect, they might be able to agree to use the recursively efficient equilibrium in the last three periods if there are no deviations in period 1, and to use the Pareto-efficient equilibrium otherwise, thus enforcing outcome (a_4, b_4) in both periods 1 and 2. (See DeMarzo 1988 and Greenberg 1988 for other discussions of equilibrium refinements as social norms.)

5.4.3 Renegotiation-Proofness in Infinitely Repeated Games

Pareto perfection and recursive efficiency for finite-horizon games are both defined using backward recursion from the terminal date. Defining renegotiation or Pareto perfection for infinite-horizon games has proved to be much more difficult, and there are currently many competing definitions. One of the earliest treatments is by Farrell and Maskin (1989), who define "weak renegotiation-proofness" for infinitely repeated games. This concept extends the "bygones are bygones" flavor of Pareto perfection by requiring that the set of renegotiation-proof equilibria at date t be independent not only of the history h^t but also of calendar time t. Weak renegotiation-proofness begins with the point of view that there is an exogenously chosen set of possible equilibrium payoffs Q that is conceivable at any t and h^t, and that each payoff in Q must require only continuation payoffs corresponding to other equilibria in Q. Formally, let $c(\sigma; h^t)$ be the continuation payoffs implied by σ given history h^t, and let $C(\sigma) = \bigcup_{t,h^t} c(\sigma; h^t)$ be the set of all continuation payoffs for strategy profile σ. Then, if $v \in Q$, there must be a perfect equilibrium σ with payoffs v such that $C(\sigma) \subseteq Q$. The set Q is said to be *weakly renegotiation-proof* (WRP) if no equilibrium payoff in Q is Pareto dominated by the payoffs of another equilibrium in Q.

	C	D
C	2,2	−1,3
D	3,−1	0,0

Figure 5.5

This definition assigns a great deal of weight to the exogenous set of "social norms" Q. This allows, for example, any static equilibrium to be weakly renegotiation-proof as a one-point set. However, in the prisoner's dilemma, the "grim" strategies of initial cooperation followed by the static equilibrium forever if someone deviates are not weakly renegotiation-proof, as the payoffs corresponding to the "cooperative phase" of the strategies Pareto dominate those of the punishment phase. That is, once the payoffs of "always cooperate" are included in the set Q of possible "agreements," the players will always renegotiate from the unending punishment back to the cooperative phase. Moreover, the strategies "perfect tit for tat," defined by "play C in the first period, and subsequently play C if last period's outcome was (C, C) or (D, D); play D if last period's outcome was (D, C) or (C, D)," are not WRP either, as in the period immediately following a unilateral deviation it would be more efficient to ignore the deviation and play (C, C). These strategies are, however, subgame perfect for discount factors near 1 with the usual payoffs, i.e., those given in figure 5.5.

Nevertheless, Farrell and Maskin (1989) and van Damme (1989) have shown that cooperation is a weakly renegotiation-proof outcome in the repeated prisoners' dilemma if the discount factor is sufficiently near to 1, and indeed the folk theorem in its renegotiation-proof version holds for this game. In particular, the strategy profile where both players use the following "penance" strategy is WRP and has efficient payoffs: "Begin in the cooperative phase where both players play C. If a single player i deviates to D, switch to the punishment phase for i. In this phase, player i plays C and the other player plays D. Play remains in this phase until the first time player i plays C, at which point play returns to the cooperative phase."

The first step in verifying that this profile is WRP is to check that it is subgame perfect. In the cooperative phase, any deviation triggers a period of punishment, which is not desirable for discount factors near 1 with payoffs as in figure 5.5. When player 1 is being punished, his payoff if he conforms is $-(1-\delta)+2\delta>0$; if he deviates, he obtains $0-\delta(1-\delta)+2\delta^2$, which is less. And when player 1 is being punished, player 2's payoff is $3(1-\delta)+2\delta$, which exceeds the payoff of 2 he receives by deviating once and then conforming. So the strategy profile is subgame perfect. Moreover, none of the three continuation payoff vectors involved is Pareto dominated by the others, so the profile is WRP.

The key to obtaining efficient WRP payoffs in the repeated prisoner's dilemma is using the profile (C, D) to punish player 1, which minmaxes player 1 while rewarding player 2. In other games, there can be a tradeoff between rewarding player 2 and punishing player 1, and this can prevent the full set of efficient individually rational payoffs from being WRP. For example, in a repeated Cournot duopoly with costless production and demand $D(p) = 2 - p$, any payoff vector of the form $(x, 1 - x)$, $x \in (0, 1)$, is feasible and individually rational, but regardless of the discount factor the only efficient WRP payoffs give each firm a payoff of at least $\frac{1}{9}$. (See Farrell and Maskin 1989 for the argument.)

Pearce (1988) and Abreu, Pearce, and Stachetti (1989) develop an alternative definition of renegotiation-proofness, to which they unfortunately give the same name. Unlike Farrell and Maskin, Pearce et al. allow some of the equilibria in $C(\sigma)$ to Pareto dominate others—they do not test for "internal" Pareto consistency; instead, they use an external test: They say that σ is renegotiation-proof unless there is a continuation payoff w in $C(\sigma)$ and another subgame-perfect equilibrium σ' such that all the continuation payoffs in $C(\sigma')$ Pareto dominate w. The idea is that the agents cannot renegotiate away from w to an alternative equilibrium that would require payoffs below w in some subgame, for fear that in that subgame the players would renegotiate back to the equilibrium with payoffs w.[19] Unlike WRP, this definition typically rules out infinite repetition of a static equilibrium. For example, in the prisoner's dilemma the infinite repetition of (D, D) is ruled out by the profile where both players play perfect "tit for tat."

Moreover, there can be nontrivial symmetric equilibria σ (equilibria where the continuation payoffs depend nontrivially on the history) that are renegotiation-proof in this sense, so that at some histories h^t all players would gain by "agreeing" to play the strategies $\sigma(\tilde{h}^t)$ corresponding to a different history. For example, the profile where both players play perfect tit for tat can be shown to be renegotiation-proof, even though it is not WRP. Pearce (1988) shows that with this definition of renegotiation-proofness, the folk theorem holds for general games. Abreu, Pearce, and Stachetti (1989) obtain an exact characterization of the symmetric renegotiation-proof equilibria in a class of games that generalizes Cournot competition.

The Farrell-Maskin definition of WRP tests only for "internal Pareto consistency," and the Abreu-Pearce-Stachetti definition tests only for external consistency; both definitions may seem weaker in some ways than Pareto perfection for finitely repeated games. One alternative would be to take a payoff that is Pareto efficient in the set of all payoffs of WRP theories

19. Pearce's definition also covers the games with imperfectly observed actions discussed in the next section. If there is a positive probability of any (finite) sequence of observations even if no player cheats, then the "subgames" (observations) that required payoffs below w have positive probability, and this definition seems better founded.

(assuming that the set of all WRP payoffs is closed). Farrell and Maskin propose several alternative definitions. A set of payoffs Q is "strongly renegotiation-proof" if it is weakly renegotiation-proof, and there is no other WRP set with a payoff that strictly Pareto dominates any of the payoffs of Q. The idea is roughly that at any time the players are able to renegotiate to a different WRP set Q' and an initial equilibrium from that theory, so that all the payoffs in Q must be immune to this sort of renegotiation. Unfortunately, such strongly renegotiation-proof payoffs need not exist, as was also observed by Bernheim and Ray (1990). Farrell and Maskin and Bernheim and Ray go on to develop more complicated solution concepts that relax strong renegotiation-proofness enough to guarantee existence.

5.5 Repeated Games with Imperfect Public Information[††]

In the repeated games considered in the last section, each player observed the actions of the others at the end of each period. In many situations of economic interest this assumption is not satisfied, because the information that players receive is only an imperfect signal of the stage-game strategies of their opponents. Although there are many ways in which the assumption of observable actions can be relaxed, economists have focused on games of *public information*: At the end of each period, all players observe a "public outcome," which is correlated with the vector of stage-game actions, and each player's realized payoff depends only on his own action and the public outcome. Thus, the actions of a player's opponents influence his payoff only through their influence on the distribution of outcomes. Games with observable actions are the special case where the public outcome consists of the realized actions themselves.

There are many examples of games in which the public outcome provides only imperfect information. Green and Porter (1984) published the first formal study of these games in the economics literature. Their model, which was intended to explain the occurrence of "price wars," was motivated in part by the work of Stigler (1964). In Stigler's model, each firm observes its own sales but not the prices or quantities of its opponents. The aggregate level of consumer demand is stochastic. Thus, a fall in a firm's sales might be due either to a fall in demand or to an unobserved price cut by an opponent. Since each firm's only information about its opponents' actions is its own level of realized sales, no firm knows what its opponents have observed, and there is no public information about the actions played.[20] In contrast, the Green-Porter model does have public information, which

20. Lehrer (1989) and Fudenberg and Levine (1990) study repeated games with imperfect private information.

makes it much easier to analyze. In that model, each firm's payoff depends on its own output and on the publicly observed market price. Firms do not observe one another's outputs, and the market price depends on an unobserved shock to demand as well as on aggregate output. Hence, an unexpectedly low market price could be due either to unexpectedly high output by an opponent or to unexpectedly low demand.

Another example of repeated games with imperfect public information is the partnership models considered by Radner (1986) and others. In these models, each player's payoffs depend on his own effort and on the publicly observed output, each player does not observe his partner's effort, and output is stochastic. Yet another example is a "noisy" prisoner's dilemma where players sometimes inadvertently choose the "wrong" action, so that the observed actions are only an imperfect signal of the intended ones. (Equivalently, each player might sometimes misperceive his opponent's action, with the payoffs a function of the perceived actions and not the intended ones.)

In the standard terminology, the above games are all examples of "repeated moral hazard." The class of games with imperfect public information can be extended to include games of "repeated adverse selection," where player i's stage-game actions a_i are maps from some private information (i.e. "types") to a space of physical actions or announcements, and all that is observed is the realized action. (The function from types to actions is not observed.) An example is Green's (1987) model of repeated insurance, in which the players' endowments are random and independent over time and between players and the stage-game strategies are maps from endowments to "announced" endowment levels. Here the reported endowments are observed, but not the maps from true endowments to reports.

5.5.1 The Model

In the *stage game*, each player $i = 1, \ldots, I$ simultaneously chooses a strategy a_i from a finite set A_i. Each action profile $a \in A = \times_i A_i$ induces a probability distribution over the publicly observed outcomes y, which lie in a finite set Y. Let $\pi_y(a)$ denote the probability of outcome y under a, and let $\pi(a)$ denote the probability distribution, which we will sometimes view as a row vector. Player i's realized payoff, $r_i(a_i, y)$, is independent of the actions of other players. (Otherwise, player i's payoff could give him *private* information about his opponents' play.) Player i's expected payoff under strategy profile a is

$$g_i(a) = \sum_y \pi_y(a) r_i(a_i, y).$$

The payoffs and distributions over outcomes corresponding to mixed strategies α are defined in the obvious way.

In the repeated game, the *public information* at the beginning of period t is

$$h^t = (y^0, y^1, \ldots, y^{t-1}).$$

Player i also has private information at time t—namely, his own past choices of actions; denote this by z_i^t. A *strategy for player i* is a sequence of maps from player i's time-t information to probability distributions over A_i; $\sigma_i^t(h^t, z_i^t)$ denotes the probability distribution chosen when player i's information is (h^t, z_i^t).

Here are some illustrations of the model:

• In a repeated game with observable actions, the set Y of outcomes is isomorphic to the set A of action profiles: $\pi_y(a) = 1$ if y is equivalent to a, and $\pi_y(a) = 0$ otherwise.

• In the Green-Porter model, $a_i \in [0, \bar{Q}]$ is firm i's output, and the outcome y is the market price. Green and Porter make the additional assumptions that the probability distribution over outcomes depends only on the sum of the firms' outputs and that every price has positive probability under every action profile.

• In the repeated partnership model, a_i is player i's effort level and y is the realized output. In the model of Radner (1986) and Radner et al. (1986), A_i is the set {work, shirk}. Closely related is the repeated principal-agent model of Radner (1981, 1985), where the principal's action is an observed monetary transfer and the agent's effort level is not observed. Here the outcome is the pair (output, transfer).

• In Green's (1987) model of repeated insurance, each period t each player i learns his current endowment θ_i^t, with the θ_i^t distributed i.i.d. according to a known distribution $P_i(\cdot)$. Here a_i is a map from the set Θ_i of all possible types to reports $\hat{\theta}_i \in \Theta_i$. (See chapter 7 for an introduction to static mechanism design.) The public outcome y^t is then the vector $\hat{\theta}^t$ of reports, which reveals neither the players' actual types nor the strategy that they used. (There are only $\prod_{i=1}^{I} (\#\Theta_i)$ outcomes, but there are $\prod_{i=1}^{I} (\#\Theta_i)^{\#\Theta_i}$ strategy profiles.) In this case the private information of player i must be extended to include the past values of his types in addition to his past actions. We will not pursue this extension here; see Fudenberg, Levine, and Maskin 1990 for details.

• In a "noisy prisoner's dilemma," the set of outcomes Y is isomorphic to the action space A, but $\pi_y(a) > 0$ even if y does not correspond to a. For example, if both players played $a_i = C$ the distribution on outcomes might be

$$\pi_{(C,C)}(C, C) = (1 - \varepsilon)^2,$$

$$\pi_{(C,D)}(C, C) = \pi_{(D,C)}(C, C) = \varepsilon(1 - \varepsilon),$$

and

$$\pi_{(D,D)}(C,C) = \varepsilon^2$$

for some $0 < \varepsilon < \frac{1}{2}$. This describes a situation where each player has probability ε of making a "mistake," and mistakes are independent. The key assumption here is that the intended actions are not observed, only the realized ones.

5.5.2 Trigger-Price Strategies

In the analysis of their oligopoly model, Green and Porter (1984) focus on equilibria in "trigger-price strategies," which generalize the trigger-strategy equilibria introduced by Friedman (1971). Suppose that the set of outcomes Y are interpreted as prices, so that $Y \subseteq \mathbb{R}$, and each firm's output a_i must lie in the interval $[0, \bar{Q}]$. Payoff functions are assumed to be symmetric and attention is restricted to equilibria where all players choose the same actions in every period—that is, $\sigma_i(h^t) = \sigma_j(h^t)$ for all t and h^t. (Thus, the equilibria are "strongly symmetric" in the sense of subsection 5.1.3.) Trigger-price-strategy profiles are indexed by three parameters, $\hat{a}$, $\hat{y}$, and $\hat{T}$. In these profiles, play can be in one of two possible "phases." In the "cooperative phase," all firms produce the same output, $\hat{a}$. Play remains in the cooperative phase as long as each period's realized price y^t is at least the "trigger price" $\hat{y}$. If $y^t < \hat{y}$, then play switches to a "punishment phase" for $\hat{T}$ periods. In this phase, the players play a static Nash equilibrium a^* in each period, regardless of the realized outcomes; after the $\hat{T}$ periods end, play returns to the cooperative phase.

If we simply take $\hat{a} = a^*$, the strategies prescribe that the static equilibrium a^* be played every period, which is clearly an equilibrium, so trigger-price equilibria exist. More generally, we can characterize the trigger-price equilibria as follows: For fixed $\hat{y}$ and $\hat{a}$, let

$$\lambda(\hat{a}) = \text{Prob}(y^t \geq \hat{y} \mid \hat{a})$$

be the probability that the outcome is at least the trigger level when players use profile $\hat{a}$. For convenience, normalize the payoff of the static equilibrium a^* to be 0. Then the (normalized) payoff if players conform to the strategies is

$$\hat{v} = (1-\delta)g(\hat{a}) + \delta\lambda(\hat{a})\hat{v} + \delta(1-\lambda(\hat{a}))\delta^{\hat{T}}\hat{v}, \tag{5.15}$$

so that

$$\hat{v} = \frac{(1-\delta)g(\hat{a})}{1-\delta\lambda(\hat{a}) - \delta^{\hat{T}+1}(1-\lambda(\hat{a}))}. \tag{5.16}$$

Note that $\hat{v} = g(\hat{a})$ if $\lambda(\hat{a}) = 1$, so that the probability of punishment is 0 so long as all players conform, or if $\hat{T} = 0$, so that "punishments" have length 0. The latter case is possible only if $\hat{a}$ is a static equilibrium, so that no punishment is needed to provide incentives. Even if $\hat{a}$ is not a static equilibrium, it might be that $\lambda(\hat{a}) = 1$, so that there is no punishment unless someone deviates; this is possible, for example, if the actions are perfectly observed. However, under the Green-Porter "full support" assumption that $\pi_y(a) > 0$ for all $y \in Y$ and all $a \in A$, the only trigger-price strategies where punishment never occurs so long as no player deviates also have the property that punishment never occurs after any sequence of outcomes. Since such strategies give players no incentive to look beyond their short-run interest, the only trigger-price equilibria where punishment never occurs are those in which there is repeated play of the static equilibria. Thus, $\lambda(\hat{a})$ will be less than 1 in equilibria that improve on the static equilibrium payoffs, and so there is a cost to imposing strong punishments for deviation. In particular, for fixed $\hat{y}$ and $\hat{a}$, the equilibrium payoffs decrease in the punishment length.

However, very long and even infinite punishments may be optimal, as by increasing the punishment length it may be possible to decrease the trigger price or increase the payoffs in the cooperative phase. The optimal trigger-price equilibria will maximize $\hat{v}$ given by equation 5.16 subject to the incentive constraint that no player gain by deviating in the cooperative phase, which is displayed in equation 5.17:

$$\begin{aligned}&(1-\delta)g(a_i,\hat{a}_{-i}) + \delta\,\lambda(a_i,\hat{a}_{-i})\hat{v} + \delta(1-\lambda(a_i,\hat{a}_{-i}))\delta^{\hat{T}}\hat{v}\\ &\quad\le (1-\delta)g(\hat{a}) + \delta\,\lambda(\hat{a})\hat{v} + \delta(1-\lambda(\hat{a}))\delta^{\hat{T}}\hat{v} \quad \text{for all } a_i. \end{aligned} \tag{5.17}$$

(No player can gain by deviating in the punishment phase, since play there is a fixed number of repetitions of a static equilibrium.)

Grouping terms together and substituting for $\hat{v}$ from equation 5.16, we get

$$\begin{aligned}&(1-\delta)[g(a_i,\hat{a}_{-i}) - g(\hat{a})]\\ &\quad\le \frac{\delta[1-\delta^{\hat{T}}][\lambda(\hat{a}) - \lambda(a_i,\hat{a}_{-i})](1-\delta)g(\hat{a})}{1-\delta\,\lambda(\hat{a}) - \delta^{\hat{T}+1}(1-\lambda(\hat{a}))}\end{aligned} \tag{5.18}$$

for all a_i.

The optimal trigger-price equilibrium (from the viewpoint of the firms) is given by the $\hat{a}$, $\hat{T}$, and $\hat{y}$ that maximize equation 5.16 subject to equation 5.18. Porter (1983b) characterizes the optimal trigger-price equilibria with a continuum of output levels and prices, and provides conditions for infinite punishments to be optimal. With a continuum of actions, the best equilibrium is better than the static one, because if the output in the cooperative phase is just a small ε below the static equilibrium levels, payoffs in the cooperative phase are greater than in the static equilibrium, while the incentive to deviate—the left-hand side of equation 5.18—is 0 to first order in ε, so that preventing deviations requires only a probability of punishment that is 0 to first order as well.

In the trigger-price equilibria, there is probability 1 that play eventually enters the punishment phase. This is loosely consistent with the idea of "price wars," but note that in equilibrium all players correctly forecast that their opponents will never deviate. Thus, the "price war" is *not* triggered by the inference that some firm chose high output in the previous period. Rather, all players correctly presume that their opponents chose the "cooperative" output last period, and that price was low because of a demand shock, but the "punishment" occurs anyway as a self-enforcing reaction to a low level of realized demand. (The solution concepts of section 5.4 were introduced in response to the concern that such punishments might not be carried out. Note that if the punishment did not occur when demand was low, players could not trust each other in the cooperative phases.)

The study of trigger-price equilibria leaves open the question of whether there are other equilibria with higher payoffs. By analogy with games with observable actions, one suspects that there may be "punishment equilibria" with payoffs lower than those in the static equilibria, and that one might in some cases be able to do better by using stronger punishments. However, this analogy is inconclusive because the punishments may be carried out even if there are no deviations. This question is one of the motivations of the Abreu-Pearce-Stachetti papers we discuss below.

5.5.3 Public Strategies and Public Equilibria

Though all the players know the public history h^t at date t, each player i also knows z_i^t, the actions he has chosen in the past. We will restrict our attention to equilibria in "public strategies," where players ignore their private information in choosing their actions.

Definition 5.2 Strategy σ_i is a *public strategy* if $\sigma_i^t(h^t, z_i^t) = \sigma_i^t(h^t, \tilde{z}_i^t)$ for all periods t, public histories h^t, and private histories z_i^t and $\tilde{z}_i^t$.

Although not all pure strategies are public strategies, it is easy to see that any payoff to a pure-strategy equilibrium is a payoff of an equilibrium in public strategies. That is, given a pure-strategy equilibrium where players' strategies may depend on their private information, we can find an equivalent equilibrium where the players' strategies depend only on their public information. The idea is that, in a pure-strategy equilibrium, each player perfectly forecasts how each opponent will play in each period — player 1 plays, say, a_1^0 in the first period, and is supposed to play $\sigma_1^1(a_1^0, y^0)$ in the second period—but since player 1's first-period play was deterministic, the conditioning of his second-period play on his first action is redundant—we could replace σ_1^1 by the public strategy $\hat{\sigma}_1^1(y^0) = \sigma_1(a_1^0, y^0)$.

When all players use public strategies, they agree about the subsequent probability distribution of actions and outcomes given any public history h^t. Thus, we can define the continuation payoffs conditional on a public history, and ask whether a profile of public strategies induces a Nash equilibrium from date t on.

Definition 5.3 A profile $\sigma_i = \{\sigma_1, \ldots, \sigma_I\}$ of repeated-game strategies is a perfect public equilibrium if

(i) each σ_i is a public strategy, and

(ii) for each date t and history h^t the strategies yield a Nash equilibrium from that date on.

Note that subgame perfection would not be restrictive in these games, since the only proper subgame is the game starting from date 0: At subsequent dates, the players need not know each other's past moves, and thus the continuation games do not emanate from a single node. However, when players use public strategies, their private information about their own past actions is irrelevant, and so perfect public equilibrium is an obvious extension of the idea of subgame perfection.

A key fact about perfect public equilibria (PPE) is that the payoffs to such equilibria are stationary—that is, the set of possible continuation payoffs of PPE starting in period t with an arbitrary public and private history is same as the set of PPE payoffs starting in period 0. (Exercise: Check this formally.) However, the sets of Nash and sequential equilibria are not, in general, stationary. (Another way of saying this is that the game lacks a "recursive structure.") Loosely speaking, the point is the following: If players 1 and 2 play a mixed strategy in the first period and their actions in the second period depend on their realized first-period action, then the actions to be played in the second period are not common knowledge. Since (in any Nash equilibrium) the first-period strategies are necessarily common knowledge, the strategic possibilities in the first and second periods are different. Exercise 5.10 develops this point further, with an example of a game which has an equilibrium that holds a player to a payoff *below* his minmax level.[21]

5.5.4 Dynamic Programming and Self-Generation

A useful tool for the analysis of perfect public equilibria is the concept of *self-generation*, introduced in Abreu, Pearce, and Stachetti 1986 and developed further in Abreu, Pearce, and Stachetti 1990. Self-generation is a sufficient condition for a set of payoffs to be supportable by perfect public equilibria. It is the multi-player generalization of the principle of optimality of discounted dynamic programming, which gives a sufficient condition for

21. The nonstationarity arises from the possibility that the players may come to have imperfectly correlated forecasts of one another's play. The same sort of imperfectly correlated forecasts arise if players observe private, correlated signals at the start of period, as in the "extensive-form correlated equilibrium" discussed in chapter 8. The set of extensive-form correlated equilibria is stationary, because the imperfect correlation that arises in period 2 from observing the public outcome in period 1 can be reproduced in period 1 with the appropriate distribution over private signals.

a vector of payoffs, one for each state, to be the maximal present values obtainable when commencing play in the corresponding state.

The key difference between self-generation in repeated games and dynamic programming is that in the former the states and the state transition function are exogenous. In repeated games, the physical environment is memoryless—the past has no physical influence on the present and the future. However, each player's strategy can depend on the history—for example, player 1's output today may depend on last period's price, and then the output that player 2 wishes to choose today might depend on last period's price as well. Thus, the control problem faced by each individual player can depend on the history, even though the physical environment does not.

Let us look at the Abreu-Pearce-Stachetti characterization of equilibrium. Recall that Y is the space of publicly observable outcomes, and let w be a function from Y into $\mathbb{R}^I$. The function w is interpreted as being the players' (normalized) continuation payoffs as a function of the realized outcome, but at this point no restrictions are made on the range of W. (Abreu, Pearce, and Stachetti use a model with a continuum of publicly observed outcomes y; we assume a finite number of outcomes for simplicity.)

Definition 5.4 The pair (α, v) is *enforceable* with respect to δ and $W \subseteq \mathbb{R}^I$ if there exists a function w: $Y \to W$ such that, for each player i,

(i) $v_i = (1 - \delta)g_i(\alpha) + \delta \sum_y \pi_y(\alpha)w_i(y)$

and

(ii) α_i solves $\max_{\alpha_i'} \left((1 - \delta)g_i(\alpha_i', \alpha_{-i}) + \sum_y \pi_y(\alpha_i', \alpha_{-i})w_i(y)\right)$.

Condition ii says that playing α_i is an optimal choice if the continuation payoffs are given by $w(\cdot)$; condition i says that when all players play α, the resulting normalized payoffs are v. Clearly, in any period t of any PPE, the actions $\sigma(h^t)$ are enforced by the equilibrium continuation payoffs; otherwise, some player could gain by a one-period deviation.

If, for some v, (α, v) is enforceable with respect to δ and W, we say that α is *enforceable* on W. If, for some α, (α, v) is enforceable with respect to δ and W, we say that v is *generated* by (δ, W). The set of all payoffs v generated by (δ, W) is denoted $\mathrm{B}(\delta, W)$.

Let $\mathrm{E}(\delta)$ denote the set of all PPE payoffs for a given discount factor. It should be clear that $\mathrm{E}(\delta) = \mathrm{B}(\delta, \mathrm{E}(\delta))$. Given any $v \in \mathrm{B}(\delta, \mathrm{E}(\delta))$, it is easy to construct a PPE with payoffs v: Choose an α and a w with range in $\mathrm{E}(\delta)$ such that w enforces (α, v), and specify that players use α in the first period and a PPE with payoffs $w(y)$ if outcome y occurs. Hence, $\mathrm{B}(\delta, \mathrm{E}(\delta)) \subseteq \mathrm{E}(\delta)$. Conversely, if $v \in \mathrm{E}(\delta)$, then no player wishes to deviate from the first-period action profile, and the continuation payoffs must (from the perfectness requirement) be in $E(\delta)$. Hence, $E(\delta) \subseteq B(\delta, E(\delta))$.

Definition 5.5 W is self-generating if $W \subseteq B(\delta, W)$.

In words, W is self-generating if the set of payoffs that can be enforced with continuation payoffs in W includes all of W. A trivial example of a self-generating set is the payoffs of a static equilibrium; static equilibria are the only one-point self-generating sets. At the other extreme, the set $E(\delta)$ of all PPE payoffs is self-generating.

Theorem 5.10[22] (Abreu, Pearce, and Stachetti 1986, 1990) If W is self-generating, then $W \subseteq E(\delta)$: All payoffs in W are PPE payoffs.

Proof Fix a $v \in W$. We will exhibit strategies for the repeated game that yield payoff v, and check that the strategies are a PPE. Since W is self-generating, $v \in B(\delta, W)$, so we have an action profile α and a map $w\colon Y \to W$ that generate payoff v. Set the period-0 strategies to be $\sigma^0 = \alpha^0$, and for each period-0 outcome y^0 set $v^1 = w^0(y^0)$. Since $v^1 \in W \subseteq B(\delta, W)$, there is an action profile $\alpha(v^1)$ and a map $w^1(y^1)\colon Y \to W$ that generates payoff v^1. Set $\sigma^1(y^0) = \alpha^1(w^0(y^0))$, and for each sequence y^0, y^1 set $v^2 = w^1(w^0(y^0))(y^1)$, and so on: The constructed strategies yield payoff v if there are no deviations, and they have been constructed so that there is no history where a player can gain by deviating once and conforming thereafter. Thus, the constructed strategies are a PPE. ■

As we remarked above, this argument is essentially that of dynamic programming, applied to a game where the physical situation is memoryless, but the past matters because it influences the opponents' play. Here, the "state" is summarized by the current target payoff v—associated with each payoff vector v we have a first-period action for each player, and a rule that specifies the continuation payoffs as a function of this period's realized outcome.

Example 5.3

To help fix ideas, here is an example of a self-generating set in a game where actions are perfectly observed, namely the prisoner's dilemma with payoffs as in figure 5.5. With observed actions, there are four outcomes y, corresponding to the four action profiles of the stage game, and the probability distribution over outcomes assigns probability 1 to the action profile that was played. Consider the two-point set $W = \{v, \hat{v}\}$, where

$$v = \begin{bmatrix} 3 - \delta & 3\delta - 1 \\ 1 + \delta\text{'} & 1 + \delta \end{bmatrix} \tag{5.19}$$

22. Abreu, Pearce, and Stachetti consider only pure-strategy equilibria, but the proof extends immediately to all PPE.

and

$$\hat{v} = \left[\frac{3\delta - 1}{1 + \delta}, \frac{3 - \delta}{1 + \delta}\right]. \tag{5.20}$$

We claim that this set is self-generating for $\delta > \frac{1}{3}$.

Given the symmetry of W, it suffices to check that payoff vector v can be enforced with continuation payoffs in W. Let the action profile α corresponding to v be (D, C), and let the continuation payoffs be $w(\text{D},\text{C}) = w(\text{C},\text{C}) = \hat{v}$ and $w(\text{D},\text{D}) = w(\text{C},\text{D}) = v$. If both players follow α, the resulting payoffs are

$$(1 - \delta)(3, -1) + \delta\hat{v} = \left[\frac{3(1 - \delta^2) + 3\delta^2 - \delta}{1 + \delta}, -\frac{(1 - \delta^2) + 3\delta - \delta^2}{1 + \delta}\right]$$

$$= v.$$

Since player 1's current action does not influence the continuation payoffs, his average payoff is maximized by playing D, as this maximizes his current payoff. If player 2 plays C as prescribed by α, his payoff is $v_2 = (3\delta - 1)/(1 + \delta)$. If he plays D, he receives payoff 0 today, and continuation payoff v_2, so that C is better than D if $\delta > \frac{1}{3}$.

Abreu, Pearce, and Stachetti (1990) prove that the set of pure-strategy equilibria is compact. There are thus best and worst equilibria. Even though those authors assume a finite number of output levels, this is not immediate, because (in contrast with our finite model) they allow a continuum of prices, so that the number of outcomes is uncountable. Furthermore, and relatedly, Abreu et al. (1986) show that any payoff to a symmetric PPE can be enforced with strategies that threaten to switch to either best or worst equilibria. There is no need for intermediate values. And, more generally, Abreu et al. (1990) show that any pure-strategy PPE payoff can be achieved with continuation values that are extremal points of the equilibrium set. Furthermore, under an additional mild condition, they show that an extremal equilibrium—an equilibrium whose payoffs are on the boundary of the feasible set—*must* have continuation payoffs that are themselves extremal equilibria.

Knowing that it is sufficient to use extremal equilibria as continuation equilibria is particularly useful for characterizing "strongly symmetric" equilibria of symmetric games—that is, equilibria where for every public history all players' actions are identical. (Strong symmetry is discussed in subsection 5.1.3; recall that the trigger-price strategies of subsection 5.5.2 are strongly symmetric.) To characterize strongly symmetric equilibria, only two numbers need be determined: the highest and the lowest strongly symmetric equilibrium payoffs, $\bar{v}$ and $\underline{v}$.

5.6 The Folk Theorem with Imperfect Public Information††

Fudenberg, Levine, and Maskin (1990) develop the dynamic-programming approach to equilibrium further and use it to prove a folk theorem for games with imperfect public information.[23] The key question in determining when the folk theorem obtains is: How much information must the public outcome reveal about the players' actions? If players receive no information at all about one another's play, the only equilibrium payoffs will be convex combinations of the payoffs to static equilibria; when actions themselves are observed, the folk theorem obtains under the mild "full-dimensionality" condition.

To begin, consider an extremal payoff v of the feasible set—that is, a point that is not a convex combination of any two other points in V. If there is an equilibrium whose payoff is close to v, it must be possible to enforce a strategy profile a with $g(a)$ close to v. When will this be the case? That is, when will there be some (not necessarily feasible) continuation payoffs that induce the players to play a? The answer is that a is enforceable unless for some player i there is an action a_i' such that

(i) $g_i(a_i', a_{-i}) > g_i(a)$

and

(ii) $\pi_{\cdot}(a_i', a_{-i}) = \pi_{\cdot}(a)$.

Condition i implies that player i prefers a_i' to a_i if the expected continuation payoffs are the same, and condition ii ensures that the two actions induce the same distributions of outcomes and thus the same distributions of continuation payoffs. It should be clear that these conditions preclude enforceability; it is also true that when the conditions fail then a is enforceable.

A slightly stronger sufficient condition for enforceability is the following *individual full-rank condition*, which implies that any two distinct mixed strategies for player i lead to different distributions over outcomes.

Definition 5.6 The *individual full-rank condition* is satisfied at profile α if for each player i the vectors $\{\pi_{\cdot}(a_i', \alpha_{-i})\}_{a_i' \in A_i}$ are linearly independent.

To see why this is called a "full-rank" condition, fix a profile α, let $\Pi_i(\alpha_{-i})$ denote the matrix whose rows are the vectors $\pi_{\cdot}(a_i', \alpha_{-i})$ corresponding to each a_i', and let $G_i(\alpha_{-i})$ denote the column vector whose elements are $[(1-\delta)/\delta]g_i(a_i', \alpha_{-i})$. Then player i has the same overall payoff

23. Their work extends an earlier result of Fudenberg and Maskin (1986b) on repeated principal-agent games. Matsushima (1989) obtains a partial folk theorem in a model with a continuum of actions on the hypothesis that for each h^t the incentive constraint can be replaced by the corresponding first-order condition.

to each action a_i' under continuation payoffs $w_i(\cdot)$ if and only if, for some constant vector k,

$$\Pi_i(\alpha_{-i}) \circ w_i = -G_i(\alpha_{-i}) + k. \tag{5.21}$$

The individual full-rank condition ensures that matrix $\Pi_i(\alpha_{-i})$ has full row rank, so that equation 5.21 can be solved for any k. (See subsection 7.6.1 for related ideas.) Note that this full-rank condition requires that there be at least as many publicly observed outcomes as there are actions for any player.

However, enforceability of all the extremal actions is not sufficient for a folk theorem in the limit of discount factors tending to 1.[24] The first counterexample was given by Radner, Myerson, and Maskin (1986) in a repeated partnership game like the following.

Example 5.4
Each period, each of two players chooses whether to work or to shirk. Each player's payoff depends on his own effort and on the publicly observed output, which they share equally. The output has only two levels, good and bad, with the probability of good equal to $\frac{9}{16}$ if both players work, $\frac{3}{8}$ if only one of them does, and $\frac{1}{4}$ if both shirk. Note that even if both players choose to work there is a positive probability of bad output. The payoffs are as follows: Working instead of shirking has a utility cost of 1; when output is good, both players receive a payment worth 4 utils; the payment in the bad output state is worth 0. (This will be the case if both players are risk neutral, the output is either 8 or 0, and the players share the output equally.)

The individual full-rank condition is satisfied at the profile where both players work, as the matrix

$$\Pi_i(\text{work}, \text{work}) = \begin{bmatrix} \frac{9}{16} & \frac{7}{16} \\ \frac{3}{8} & \frac{5}{8} \end{bmatrix}$$

is nonsingular. Thus, both players can be induced to work by the appropriate choice of continuation payoffs. And the profile where both players work yields $(\frac{5}{4}, \frac{5}{4})$, which is the highest feasible symmetric payoff. However, regardless of δ, the sum of the equilibrium payoffs is bounded by 2. The intuition for why efficiency cannot even be approximated in this model is that, in order to provide incentives for both players to work, both players' continuation payoffs must be higher after high output than after low. Loosely speaking, this means that both players must be "punished" when a bad outcome is observed. As long as the bad outcome has positive

24. Rubinstein (1979a), Rubinstein and Yaari (1983), and Radner (1986) obtain Nash-threats folk theorems using time-average payoffs in examples of games that do not meet the stronger information conditions we develop below. The literature suggests that individual full rank suffices for the full folk theorem with time-average payoffs, but we are not aware of a formal proof.

probability when both players work, there must be a positive probability of this "mutual punishment," and since mutual punishments are inefficient, the set of equilibrium payoffs is bounded away from efficiency.

Although the bound holds for all Nash equilibria of the game, it is easiest to obtain for the symmetric pure-strategy equilibria. Let v^* be the highest payoff in any pure-strategy symmetric equilibrium. Since the set of these equilibria is stationary, the payoff in the first period of an equilibrium with payoffs v^* must be at least v^*; thus, if $2v^*$ is greater than 2, in any equilibrium with payoffs v^* both players must work in the first period. If v_g is the (symmetric) continuation payoff after the good output and v_b the continuation payoff after bad, incentive compatibility requires that

$$(1-\delta)[(\tfrac{9}{16}\cdot 4 + \tfrac{7}{16}\cdot 0) - 1] + \delta[\tfrac{9}{16}v_g + \tfrac{7}{16}v_b]$$
$$\geq (1-\delta)[\tfrac{3}{8}\cdot 4 + \tfrac{5}{8}\cdot 0] + \delta[\tfrac{3}{8}v_g + \tfrac{5}{8}v_b],$$

or

$$v_g - v_b \geq [(1-\delta)/\delta]\tfrac{4}{3}.$$

Since $v_g \leq v^*$, we conclude that if v^* is close to $\frac{5}{4}$ then

$$v^* \leq (1-\delta)\tfrac{5}{4} + \delta[\tfrac{9}{16}v^* + \tfrac{7}{16}\{v^* - [(1-\delta)/\delta]\tfrac{4}{3}\}],$$

so $(1-\delta)v^* \leq (1-\delta)\frac{8}{12}$—a contradiction.

Here, even though the required difference between v_g and v_b goes to 0 as δ goes to 1, the normalized present value of the efficiency loss remains nonnegligible.

In other cases, though, it is possible to provide all players with incentives to take the desired action while incurring a minimal loss of efficiency. In this case one can show that the efficiency loss required to provide incentives becomes negligible as $\delta \to 1$. When can the continuation payoffs be chosen in this way? *A sufficient condition is that the distributions over outcomes induced by different players' deviations be distinct.* This is made precise in the following definition and lemma.

Definition 5.7 The *pairwise full-rank condition* is satisfied at action α for players i and j if the $(|A_i| + |A_j|)$ vectors

$$\{\pi_{\cdot}(a_i', \alpha_{-i})_{a_i' \in A_i}, \pi_{\cdot}(a_j', \alpha_{-j})_{a_j' \in A_j}\}$$

admit only one linear dependency.

This condition implies that the matrix $\Pi_{ij}(\alpha)$ formed by stacking the matrix $\Pi_i(\alpha_{-i})$ on top of the matrix $\Pi_j(\alpha_{-j})$ has maximal rank. This matrix does not have *full* row rank, as it necessarily admits at least one linear dependency. This is easiest to see in the case where all players use their first pure strategy, so the first rows of Π_i and Π_j are identical. More generally, the rows of Π_{ij} satisfy the following equality:

$$\pi_{.}(\alpha) = \sum_{a_i \in A_i} \alpha_i(a_i)\pi_{.}(a_i, \alpha_{-i}) = \sum_{a_j \in A_j} \alpha_j(a_j)\pi_{.}(a_j, \alpha_{-j}). \tag{5.22}$$

If profile α satisfies pairwise full rank, not only can it be enforced, but the continuation payoffs can be chosen to satisfy the additional linear identity $\beta_1 w_1(y) + \beta_2 w_2(y) = k$ for any nonzero β_1 and β_2. That is, player i's continuation payoff can be exchanged for player j's at rate $-\beta_2/\beta_1$, so it is as if utility were transferable between the players. In this case, we can arrange the continuation payoffs so that when player i is punished, player j is rewarded, and conversely. Moreover, when profile α is efficient, the rate of exchange can be taken to be equal to the tangent to the efficient frontier at profile α, which is the key to providing incentives in an efficient way.

Under pairwise full rank, a deviation by player i leads to a distribution over outcomes that is different from that induced by any deviation by player j.

Note that this condition requires that the number of outcomes be at least $|A_i| + |A_j| - 1$. In example 5.4, there are two actions per player, and only two outcomes, so that pairwise full rank cannot be satisfied at any action profile. This is why shirking by player 1 could not be distinguished, even statistically, from shirking by player 2.

Even if the number of outcomes is large enough to permit pairwise full rank to be satisfied, the condition can still fail at some profiles. In particular, regardless of the number of outcomes, pairwise full rank fails at symmetric profiles in games such as the Green-Porter oligopoly or the partnership of example 5.4, where the distribution of outcomes depends only on the sum of the individual player's actions. For example, regardless of the number of outcomes in example 5.4, at a profile where both players work we see that

$$\Pi_{12}(\text{work}, \text{work}) = \begin{bmatrix} \pi_{.}(\text{work}, \text{work}) \\ \pi_{.}(\text{shirk}, \text{work}) \\ \pi_{.}(\text{work}, \text{work}) \\ \pi_{.}(\text{work}, \text{shirk}) \end{bmatrix},$$

and since $\pi_{.}(\text{shirk}, \text{work}) = \pi_{.}(\text{work}, \text{shirk})$ this matrix only has rank 2, instead of the rank 3 that pairwise full rank requires.

However, if there are more than two outcomes, the profile where player 1 works and player 2 shirks does satisfy pairwise full rank for generic probability distributions on outcomes. For example, suppose there are three outcomes, y_1, y_2, and y_3, and that

$$\pi_{.}(\text{work}, \text{work}) = (\tfrac{1}{2}, \tfrac{3}{8}, \tfrac{1}{8}),$$

$$\pi_{.}(\text{work}, \text{shirk}) = \pi_{.}(\text{shirk}, \text{work}) = (\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{4}),$$

and

$$\pi_{\cdot}(\text{shirk}, \text{shirk}) = (\tfrac{1}{8}, \tfrac{3}{8}, \tfrac{1}{2}).$$

Then

$$\Pi_{12}(\text{work}, \text{shirk}) = \begin{bmatrix} (\frac{1}{4}, \frac{1}{2}, \frac{1}{4}) \\ (\frac{1}{8}, \frac{3}{8}, \frac{1}{2}) \\ (\frac{1}{2}, \frac{3}{8}, \frac{1}{8}) \\ (\frac{1}{4}, \frac{1}{2}, \frac{1}{4}) \end{bmatrix},$$

which has rank 3. Moreover, as observed by Legros (1988), any profile where player 1 works and player 2 shirks with positive probability also satisfies pairwise full rank, since the profile where player 1 shirks and player 2 uses his mixed strategy induces a different distribution than the profile where player 1 works and player 2 shirks. Legros' observation is generalized in the following lemma.

Lemma 5.1 If for each pair of players $i \neq j$ there is an $\alpha^{i,j}$ that satisfies pairwise full rank for players i and j, then there is an open dense set of α's that satisfy pairwise full rank for all pairs of players.

Theorem 5.11 (Fudenberg, Levine, and Maskin 1990) If (i) the individual full-rank condition is satisfied at every pure strategy a, (ii) for each pair i, j of players there is a profile that satisfies pairwise full rank for i and j, and (iii) the feasible set V has dimension equal to the number of players, then for closed set W in the relative interior of V there is a $\underline{\delta}$ such that, for all $\delta > \underline{\delta}$, $W \subseteq \mathrm{E}(\delta)$.

Outline of Proof Approximate the set of feasible individually rational payoffs by a smooth convex set W. Condition i implies that the minmax profile against player i and the *best* profile for player i can both be enforced on hyperplanes where player i's payoff is constant. Condition ii implies that almost all profiles can be enforced on hyperplanes where no player's payoff is held constant. Combining these two observations, for any point w on the boundary of W there is an action profile α with $g(\alpha)$ weakly separated from W by the tangent plane H at w, *and* such that profile α can be enforced with continuation payoffs on any linear translate $H + v$ of H. If we choose δ to be close to 1, the required variation in $w(y)$ is small enough that profile α can be enforced with continuation payoffs contained in a translate of H very close to the boundary of W. Intuitively, the "efficiency loss" (relative to W) required to provide incentives becomes negligible, since the smooth set W is approximately (i.e., to first order) linear.

In a symmetric game, the theorem asserts that there are equilibria with payoffs arbitrarily close to the highest symmetric payoff. This is so even though in such games the highest payoff in a *symmetric equilibrium* may be bounded away from efficiency. The point is that the information revealed at symmetric action profiles can be poor (i.e., fail to satisfy pairwise full rank) even though many nearby almost-symmetric strategy profiles do generate "enough" information.[25]

5.7 Changing the Information Structure with the Time Period†††

The folk theorem looks at a set of equilibrium payoffs as $\delta \to 1$, holding $\pi_y(a)$ constant. As we saw, whether the folk theorem holds depends on the amount of information the public outcome y reveals. The interpretation of the result is therefore that almost all feasible, individually rational payoffs are equilibrium payoffs when δ is large in comparison with the information revealed by the outcome. Abreu, Pearce, and Milgrom (1990) show that the folk theorem need not hold if one interprets $\delta \to 1$ as the result of the interval between periods converging to 0, and if the information revealed by y deteriorates as the time interval shrinks. Why might this be the case? In games with observed actions, the public outcome is perfectly informative, and there is no reason to expect the information to change as the time period shrinks. In these games, then, we can interpret $\delta \to 1$ as a situation of either very little time preference or very short time periods. However, if players observe only imperfect signals of one another's actions, it is plausible that the quality of their information depends on the length of each observation period. Thus, one cannot interpret the case of $\delta \cong 1$, with $\pi_y(a)$ fixed, as the study of what would occur if the time period became very short.

Abreu, Pearce, and Milgrom (APM) investigate the effects of changing the time period and the associated information structure in two different examples. We will focus on a variant of their first example, a model of a repeated partnership game. We begin as usual by describing the stage game, which in the APM model is a continuous-time game of length τ. The interpretation is that players lock in their actions at the start of the stage, and at the end of the stage the outcome and the payoffs are revealed. As in example 5.4, each player has two choices: work and shirk. Payoffs are chosen so that shirk is a dominant strategy in the stage game, and so that shirk is the minmax strategy. As in the example, the stage game has the structure of the prisoner's dilemma: "Both shirk" is a Nash equilibrium in dominant strategies, and this equilibrium gives the players their minmax values. Payoffs are normalized so that this minmax payoff is 0, the (ex-

25. Fudenberg, Levine, and Maskin go on to develop a Nash-threats folk theorem for games where there are too few outcomes for even the individual full-rank condition to hold but where the information revealed by the outcomes has a "product structure," meaning that $y = (y_1, \ldots, y_I)$, and each player i's action influences only his "own" outcome y_i. This is the case in Green's (1987) model of repeated adverse selection, where the actions are reports of the players' types.

pected) payoffs if both players work are (c, c), and the payoff to shirking when the opponent works is $c + g$. (These are the expected payoffs, where the expectation is taken with respect to the corresponding distribution of output.) The difference between the APM stage game and example 5.4 is that, instead of there being only two outcomes each period (namely high and low output), the outcome is the number of "successes" in the period, which is distributed as a Poisson variable whose intensity is λ if both players work and μ if one of them shirks, with $\lambda > \mu$. Thus, if the time period is short, it is unlikely that there will be more than one success, and the probability of one success in a period of length dt is proportional to dt. This might correspond to a situation where the workers are trying to invent new products.[26]

In the repeated game $G(\tau, r)$, the discount factor δ is $\exp(-r\tau)$, and the public information is simply the number of "successes" each period.

From our earlier discussion of repeated partnership games, we can see that the folk theorem does not apply to the symmetric equilibria of this game, as deviations by the two players are indistinguishable when both use the same strategy. Thus, both players must be punished simultaneously, which causes efficiency losses, and so the set of symmetric equilibria is bounded away from efficiency even as $r \to 0$. However, when r is small, we would expect there to be symmetric equilibria with higher payoffs than the static equilibrium. Even this limited conclusion does not hold in the limit $\tau \to 0$, as the information revealed by the outcomes may "deteriorate" quickly enough to outweigh the effects of a larger discount factor. APM compute the highest symmetric-equilibrium payoffs of the game for small τ, and consider how the payoff of the best symmetric equilibrium varies with τ. To do so they consider a Taylor-series approximation of the game, neglecting terms smaller than τ^2.

This Taylor-series approach is necessary to discover if it is best to increase or decrease τ when τ is small. We will content ourselves with making the simpler point that sending τ to 0, with the corresponding changes in the information structure, has very different effects than sending r to 0, holding the information structure fixed. To this end we simplify the APM model by assuming that there are only two possible outcomes in each period: For $\theta = \lambda, \mu$, there is probability $\exp(-\theta\tau)$ that no events occur, and probability $1 - \exp(-\theta\tau)$ that exactly one event occurs. This simplifies the Poisson distribution by identifying all the events with one or more outcomes; it is a good approximation to the Poisson distribution when periods are short.

Let us consider when the best symmetric pure-strategy equilibrium can have payoff v^* that strictly exceeds 0. (APM show that this maximum is

26. Abreu, Pearce, and Milgrom also consider the case of "bad news," where *low* output is a Poisson event with intensity λ if both players work and μ if one shirks, with $\lambda < \mu$. Here the Poisson event corresponds to "accidents" that are made less likely if both players work.

attained; the following arguments would extend with the addition of a few epsilons if v^* were a supremum rather than a maximum.) Since APM allow public randomizations, it is immediate that an equilibrium with payoffs (v^*, v^*) can be constructed using continuation payoffs which are lotteries between the best continuation payoff of v^* and the worst continuation payoff of 0, as any continuation payoff between these values can be obtained by a public randomization between them. Thus, fix an equilibrium that attains v^*, and let the continuation payoff for each player be the lottery $[(1 - \alpha(0))v^*, \alpha(0)\cdot 0]$ if the first-period outcome is 0, and the lottery $[(1 - \alpha(1))v^*, \alpha(1)\cdot 0]$ if the first-period outcome is 1. One can show that if v^* is greater than 0, the strategies that attain (or closely approximate) v^* must have both players working in the first period. Thus, in order for v^* to exceed 0 there must exist probabilities $\alpha(0)$ and $\alpha(1)$ such that both agents are induced to work, and such that if both players do work the resulting normalized present values are v^*.

Writing out these two equations, we have

$$(1 - e^{-r\tau})g \leq e^{-r\tau}\cdot(e^{-\mu\tau} - e^{-\lambda\tau})\cdot[\alpha(0) - \alpha(1)]v^* \tag{5.24}$$

and

$$v^* = (1 - e^{-r\tau})c + e^{-r\tau}[(1 - \alpha(0)e^{-\lambda\tau}) - \alpha(1)(1 - e^{-\lambda\tau})]v^*. \tag{5.25}$$

Solving equation 5.25 for v^* yields

$$v^* = \frac{(1 - e^{-r\tau})c}{1 - e^{-r\tau}\{1 - \alpha(1) - e^{-\lambda\tau}[\alpha(0) - \alpha(1)]\}}. \tag{5.26}$$

Intuitively, one would expect that in the best equilibrium players would not be "punished" if a success occurs, so that $\alpha(1) = 0$. This can be checked by inspecting the above equations: When $\lambda > \mu$, setting $\alpha(1) = 0$ makes equation 5.24 more likely to be satisfied, and increases the equilibrium payoff by decreasing the probability of switching to the punishment state.

Now we ask when the above system has a solution with $\alpha(1) = 0$ and $\alpha(0) \leq 1$. Algebraic manipulation shows that this is possible only if

$$c[e^{-\mu\tau} - e^{-\lambda\tau}] \geq g[e^{r\tau} - 1 + e^{-\lambda\tau}] \geq ge^{-\lambda\tau}. \tag{5.27}$$

Thus, a necessary condition for $v^* > 0$, regardless of the rate of interest, is that

$$\frac{g}{c} \leq \frac{e^{-\mu\tau} - e^{-\lambda\tau}}{e^{-\lambda\tau}} = e^{-(\mu-\lambda)\tau} - 1, \tag{5.28}$$

which says that the likelihood ratio $L(\tau) = e^{-(\mu-\lambda)\tau}$ associated with the event "no successes" should be sufficiently large. However, as τ converges to 0, the likelihood ratio $L(\tau)$ converges to 1: Since it is almost certain that there will be no successes, the information provided by the publicly observed outcome is too poor for there to be an equilibrium that improves on the minmax values.

On the topic of changing the information structure, we should also mention Kandori (1989a), who studies how the set of equilibrium payoffs changes when the public outcome becomes a less informative signal of the players' actions. One probability distribution is a "garbling" of another (Blackwell and Girshik 1954) if it can be obtained from the first one by adding noise. Kandori shows that if the information becomes worse in the sense of garbling, the set of equilibrium payoffs becomes strictly smaller. That the set cannot grow larger is fairly clear, and is obvious in the presence of public randomizations, as the public randomizing device can be used to create a garbling of the original signal; the interesting conclusion is that the set must become strictly smaller.

Exercises

Exercise 5.1* Compute the set of feasible payoffs in the "battle of the sexes" stage game as shown in figure 1.10a. What is the highest feasible symmetric payoff? Let $\delta = \frac{9}{10}$, and find a deterministic strategy profile for the repeated game with payoffs $(\frac{3}{2}, \frac{3}{2})$.

Exercise 5.2* Consider the infinitely repeated play of a finite stage game $(\mathscr{I}, A, g)$. Given $\varepsilon > 0$, show that there exists a $\underline{\delta} > 0$ such that for all $\delta \in [0, \underline{\delta}]$ every Nash equilibrium σ has the property that, at all histories h^t with positive probability under σ, $\sigma(h^t)$ must be within ε of one of the Nash equilibria of the stage game. Give an example to show that the conclusion need not hold for all subgames. Can the equilibrium to be played in period t vary with the history h^t? Why or why not?

Exercise 5.3** Prove that the set of continuation payoff vectors corresponding to all Nash equilibria is the same in every proper subgame of a repeated game. The idea of the proof is to show more strongly that every proper subgame is strategically isomorphic, i.e., there is a one-to-one correspondence between the strategy spaces that preserves the payoffs. The simplest example is the map between the whole game and a subgame: To map a strategy s for the whole game to its equivalent in the subgame starting at h^t, set $\hat{s}(h^t) = s(h^0)$, $\hat{s}(h^t, a^t) = s(a^t)$, $\hat{s}(h^t, a^t, a^{t+1}) = s(a^t, a^{t+1})$, and so on, so that $\hat{s}$ treats period $t + \tau$ just as s treats period τ. Conversely, given a strategy profile s and a subgame h^t, the equivalent strategy for the whole game is $\hat{s}(h^0) = s(h^t)$, $\hat{s}(a^0) = s(h^t, a^0)$, and so on. Use these maps between the subgames to argue that if a profile for the whole game is a Nash equilibrium it must be a Nash equilibrium in the subgame, and conversely.

Exercise 5.4* Consider a finite symmetric repeated game, and assume there is a symmetric mutual minmax profile m^* in pure strategies, i.e., a pure-strategy profile m such that $\max_{a_i} g(a_i, m^*_{-i}) \leq \underline{v}$. Show that, if public randomizations are available, for sufficiently large discount factors the worst strongly symmetric equilibrium payoff e_* can be attained with strategies that have two phases: In phase A, players play m^*. If players conform in phase A, play switches to phase B with a probability specified by the equilibrium strategies; if there are any deviations, play remains in phase A with probability 1. In phase B, play follows strategies that yield the highest equilibrium payoff.

Exercise 5.5* Consider the two-player game illustrated in figure 5.6. In the first period, players 1 and 2 simultaneously choose U1 or D1 (player 1) and L1 or R1 (player 2); these choices are revealed at the end of period 1 with payoffs as in the left-hand matrix. In period 2, players choose U2 or D2 and L2 or D2, with payoff as in the matrix on the right. Each player's objective is to maximize the average of his per-period payoff.

(a) Find the subgame-perfect equilibria of this game, and compute the convex hull of the associated payoffs.

(b) Now suppose that the players can jointly observe the outcome y_1 of a public randomizing device before choosing their first-period actions, where y_1 has a uniform distribution on the unit interval. Find the set of subgame-perfect equilibria, and compare the resulting payoffs against the answer to part a of this exercise.

(c) Suppose that the players jointly observe y_1 at the beginning of period 1 and y_2 at the beginning of period 2, with y_1 and y_2 being independent draws from a uniform distribution on the unit interval. Again, find the subgame-perfect equilibrium payoffs.

(d) Relate your answers to parts a–c to the role of public randomizations in the proof of the Folk Theorem.

Exercise 5.6* Under the assumptions used by Benoit and Krishna (1985) for pure-strategy equilibria, try to characterize the limit as the horizon $T \to \infty$ of the set of payoffs of all subgame-perfect equilibria of a T-period finitely repeated game.

	L1	R1
U1	2,2	−1,3
D1	3,−1	0,0

First-period payoffs

	L2	R2
U2	6,4	3,3
D2	3,3	4,6

Second-period payoffs

Figure 5.6

Exercise 5.7** Consider a sequence of randomly matched players, with the information structure of Rosenthal (1979), who play the prisoner's dilemma with payoffs as in figure 5.5. Show that unless the discount factor equals $\frac{1}{3}$, the only Markovian equilibrium where all players use the same pure strategy is for all players to always cheat.

Exercise 5.8**

(a) In a repeated game, show that if for each player there is a subgame-perfect equilibrium where that player's payoff is his minmax value, then any payoff of a Nash equilibrium is also the payoff of a subgame-perfect equilibrium.

(b) Suppose that, for each player i and each $j \neq i$, $g_j(m^i) > \underline{v}_j$. Show that the sets of Nash-equilibrium and perfect-equilibrium payoffs are identical for sufficiently large discount factors. Give an economic example where the condition is plausible, and an example where it is not. Show that the sets can differ for small discount factors.

(c) Suppose that the minmax profile is in pure strategies, that the vector where all players simultaneously receive their minmax payoff is in the interior of the feasible set, and that for each player i there is an $\hat{a}_i$ such that $g_i(\hat{a}_i, m^i_{-i}) < \underline{v}_i$. Show that the sets of Nash-equilibrium and perfect-equilibrium payoffs are identical for large enough discount factors. Give an example of a game where the feasible set has full dimension yet the inferiority condition used here does not apply. (Answers are given in Fudenberg and Maskin 1990b.)

Exercise 5.9* Consider infinitely repeated play of the stage game of figure 5.7.

(a) What is the highest perfect equilibrium for player 1 if both sides are long-run players?

(b) If player 1 could publicly commit to always play the same mixed strategy α_1, what α_1 would he choose? What would his payoff be?

(c) Show that when the player 2's are an infinite sequence of short-run players, the highest payoff for player 1 in any Nash equilibrium is 2. To do this, proceed as follows.

• Let $v^*(\delta)$ be the supremum of player 1's payoff in any Nash equilibrium when his discount factor is δ, and suppose $v^*(\delta) > 2$. Let $\varepsilon = (1-\delta)(v^*(\delta)-2)/2$, and choose an equilibrium σ where player 1's equilibrium payoff $v(\delta)$ is at least $v^*(\delta) - \varepsilon$. Show that under profile σ player 1's

	L	M	R
U	6,0	-1,-100	0,1
D	2,2	0,3	1,1

Figure 5.7

expected payoff in the first period must be greater than 2. (Hint: Player 1's continuation payoff from the second period on cannot exceed $v^*(\delta)$.)

• Show that under σ player 2 must play L with positive probability in the first period, and thus that player 1 must play D with positive probability in the first period.

• Conclude that $v^*(\delta) - \varepsilon \leq v(\delta) \leq 2(1 - \delta) + \delta v^*(\delta)$, so that $v^*(\delta) \leq 2$.

Exercise 5.10** Consider the following three-person stage game: Players 1 and 2 choose pairs, the first element being Up (U) or Down (D) and the second being Heads (H) or Tails (T). Player 3 chooses Right or Left. Players 1 and 2 receive 0 regardless of what happens. Player 3's payoff is -1 if 1 and 2 both chose Up and he went Right, or if 1 and 2 both chose Down and he went Left; otherwise he gets 0. Notice that the choice of H or T is irrelevant to the payoffs.

Suppose player 3's choices are observable, that whether 1 and 2 play Up or Down is observable, but that the public information y^t about their choice of H or T is only the total number of H chosen by both players. Thus, if y^t is 2 or 0, it reveals the actions of players 1 and 2. If $y^t = 1$, then the actions of players 1 and 2 are common knowledge for players 1 and 2 (since it is common knowledge that they each know their own action), but player 3 does not know which player played H.

(a) Show that player 3's minmax payoff is $-\frac{1}{4}$.

(b) Now consider the repeated version of this game. Construct a Nash equilibrium where players 1 and 2 use the following strategies, e.g., for player 1: "Randomize $\frac{1}{2}$-$\frac{1}{2}$ between H and T in every period. If $y^{t-1} = 1$ and player 2 played H, play D; if $y^{t-1} = 1$ and player 2 played T, then play U. If $y^{t-1} = 0$ or 2, randomize $\frac{1}{2}$-$\frac{1}{2}$ between U and D." Show that player 3's equilibrium payoff is below his minmax value. Explain.

References

Abreu, D. 1986. Extremal equilibria of oligopolistic supergames. *Journal of Economic Theory* 39: 191–228.

Abreu, D. 1988. Towards a theory of discounted repeated games. *Econometrica* 56: 383–396.

Abreu, D., and D. Dutta. 1990. in preparation.

Abreu, D., D. Pearce, and P. Milgrom. 1990. Information and timing in repeated partnerships. *Econometrica*, forthcoming.

Abreu, D., D. Pearce, and E. Stachetti. 1986. Optimal cartel equilibrium with imperfect monitoring. *Journal of Economic Theory* 39: 251–269.

Abreu, D., D. Pearce, and E. Stachetti. 1989. Renegotiation and symmetry in repeated games. Mimeo, Harvard University.

Abreu, D., D. Pearce, and E. Stachetti. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58: 1041–1064.

Aumann, R. 1986. Repeated games. In *Issues in Contemporary Microeconomics*, ed. G. Feiwel. Macmillan.

Aumann, R. 1989. Survey of repeated games. In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern.* Manheim: Bibliographisches Institut.

Aumann, R. 1990. Communication need not lead to Nash equilibrium. Mimeo.

Aumann, R., and L. Shapley. 1976. Long-term competition—a game theoretic analysis. Mimeo.

Benoit, J. P., and V. Krishna. 1985. Finitely repeated games. *Econometrica* 53: 890–904.

Benoit, J. P., and V. Krishna. 1987. Nash equilibria of finitely repeated games. *International Journal of Game Theory* 16.

Benoit, J. P., and V. Krishna. 1988. Renegotiation in finitely repeated games. Discussion Paper 89-004, Harvard University.

Bergin, J., and B. MacLeod. 1989. Efficiency and renegotiation in repeated games. Mimeo, Queen's University.

Bernheim, B. D., B. Peleg, and M. Whinston. 1987. Coalition-proof Nash equilibria. I: Concepts. *Journal of Economic Theory* 42: 1–12.

Bernheim, B. D., and D. Ray. 1989. Collective dynamic consistency in repeated games. *Games and Economic Behavior* 1: 295–326.

Blackwell, D. and M. Girshik. 1954. *Theory of Games and Statistical Decisions.* Wiley.

Chamberlin, E. 1929. Duopoly: Value where sellers are few. *Quarterly Journal of Economics* 43: 63–100.

Crémer, J. 1986. Cooperation in ongoing organizations. *Quarterly Journal of Economics* 101: 33–49.

DeMarzo, P. M. 1988. Coalitions and sustainable social norms in repeated games. IMSSS Discussion Paper 529, Stanford University.

Dybvig, P., and C. Spatt. 1980. Does it pay to maintain a reputation? Mimeo, Carnegie-Mellon University.

Ellison, G. 1991. Cooperation in random matching games. Mimeo.

Farrell, J., and E. Maskin. 1989. Renegotiation in repeated games. *Games and Economic Behavior* 1: 327–360.

Fisher, I. 1898. Cournot and mathematical economics. *Quarterly Journal of Economics* 12: 1–26.

Forges, F. 1986. An approach to communications equilibrium. *Econometrica* 54: 1375–1386.

Frayssé, J., and M. Moreaux. 1985. Collusive equilibria in oligopolies with long but finite lives. *European Economic Review* 27: 45–55.

Friedman, J. 1971. A noncooperative equilibrium for supergames. *Review of Economic Studies* 38: 1–12.

Friedman, J. 1985. Trigger strategy equilibria in finite horizon supergames. Mimeo.

Fudenberg, D., D. Kreps, and E. Maskin. 1990. Repeated games with long-run and short-run players. *Review of Economic Studies* 57: 555–574.

Fudenberg, D., and D. Levine. 1983. Subgame-perfect equilibria of finite and infinite horizon games. *Journal of Economic Theory* 31: 227–256.

Fudenberg, D., and D. Levine. 1990a. Approximate equilibria in repeated games with imperfect private information. *Journal of Economic Theory*, forthcoming.

Fudenberg, D., and D. Levine. 1990b. Efficiency and observability in games with long-run and short-run players. Mimeo.

Fudenberg, D., D. Levine, and E. Maskin. 1990. The folk theorem in repeated games with imperfect public information. Mimeo, Massachusetts Institute of Technology.

Fudenberg, D., and E. Maskin. 1986a. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54: 533–556.

Fudenberg, D., and E. Maskin. 1986b. Discounted repeated games with one-sided moral hazard. Mimeo.

Fudenberg, D., and E. Maskin. 1990a. On the dispensability of public randomizations in discounted repeated games. *Journal of Economic Theory*. Forthcoming.

Fudenberg, D., and E. Maskin. 1990b. Nash and perfect equilibria of discounted repeated games. *Journal of Economic Theory* 50: 194–206.

Green, E. 1987. Lending and the smoothing of uninsurable income. In *Contractual Arrangements for Intertemporal Trade*, ed. E. Prescott and N. Wallace. University of Minnesota Press.

Green, E., and R. Porter. 1984. Noncooperative collusion under imperfect price information. *Econometrica* 52: 87–100.

Greenberg, J. 1988. The theory of social situations. Mimeo, Haifa University.

Greif, A. 1989. Reputation and coalitions in medieval trade: Evidence from the geniza documents. Mimeo.

Kandori, M. 1989a. Monotonicity of equilibrium payoff sets with respect to obervability in repeated games with imperfect monitoring. Mimeo.

Kandori, M. 1989b. Social norms and community enforcement. Mimeo.

Kandori, M. 1989c. Repeated games played by overlapping generations of players. Mimeo.

Klein, B., and K. Leffler. 1981. The role of market forces in assuring contractual performance. *Journal of Political Economy* 81: 615–641.

Kreps, D. 1986. Corporate culture and economic theory. In *Technological Innovation and Business Strategy*, ed. M. Tsuchiya. Nippon Keizai Shimbunsha Press (in Japanese). Also (in English) in *Rational Perspectives on Political Science*, ed. J. Alt and K. Shepsle. Harvard University Press, 1990.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson. 1982. Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory* 27: 245–252.

Lambson, V. E. 1987. Optimal penal codes in price-setting supergames with capacity constraints. *Review of Economic Studies* 54: 385–397.

Legros, P. 1988. Sustainability in partnerships. Mimeo, California Institute of Technology.

Lehrer, E. 1988. Two player repeated games with non-observable actions and observable payoffs. Mimeo.

Lehrer, E. 1989. Lower equilibrium payoffs in two-player repeated games with non-obervable actions. *International Journal of Game Theory* 18.

Macaulay, S. 1963. Non-contractual relations in business: A preliminary study. *American Sociological Review* 28: 55–67.

Matsushima, H. 1989. Efficiency in repeated games with imperfect monitoring. *Journal of Economic Theory* 48: 428–442.

Mertens, J.-F. 1987. Repeated games. In *Proceedings of the International Congress of Mathematicians 1986*.

Mertens, J.-F., S. Sorin, and S. Zamir. 1990. Repeated games. Manuscript.

Milgrom, P., D. North, and B. Weingast. 1989. The role of law merchants in the revival of trade: A theoretical analysis. Mimeo.

Munkres, J. 1975. *Topology: A First Course*. Prentice-Hall.

Myerson, R. 1986. Multistage games with communication. *Econometrica* 54: 323–358.

Neyman, J. 1989. Counterexamples with almost common knowledge. Mimeo.

Oye, K., ed. 1986. *Cooperation under Anarchy*. Princeton University Press.

Pearce, D. 1988. Renegotiation-proof equilibria: Collective rationality and intertemporal cooperation. Mimeo, Yale University.

Porter, R. 1983a. A study of cartel stability: The joint economic committee, 1880–1886. *Bell Journal of Economics* 14: 301–314.

Porter, R. 1983b. Optimal cartel trigger-price strategies. *Journal of Economic Theory* 29: 313–338.

Radner, R. 1980. Collusive behavior in non-cooperative epsilon equilibria of oligopolies with long but finite lives. *Journal of Economic Theory* 22: 121–157.