5 Reputation and Repeated Games with Symmetric Information

23 November 2005. Eric Rasmusen, Erasmuse@indiana.edu. Http://www.rasmusen.org.

5.1 Finitely Repeated Games and the Chainstore Paradox

Chapter 4 showed how to refine the concept of Nash equilibrium to find sensible equilibria in games with moves in sequence over time, so-called dynamic games. An important class of dynamic games is repeated games, in which players repeatedly make the same decision in the same environment. Chapter 5 will look at such games, in which the rules of the game remain unchanged with each repetition and all that changes is the "history" which grows as time passes, and, if the number of repetitions is finite, the approach of the end of the game. It is also possible for asymmetry of information to change over time in a repeated game since players' moves may convey their private information, but Chapter 5 will confine itself to games of symmetric information.

Section 5.1 will show the perverse unimportance of repetition for the games of Entry Deterrence and The Prisoner's Dilemma, a phenomenon known as the Chainstore Paradox. Neither discounting, probabilistic end dates, infinite repetitions, nor precommitment are satisfactory escapes from the Chainstore Paradox. This is summarized in the Folk Theorem of Section 5.2. Section 5.2 will also discuss strategies which punish players who fail to cooperate in a repeated game— strategies such as the Grim Strategy, Tit-for- Tat, and Minimax. Section 5.3 builds a framework for reputation models based on The Prisoner's Dilemma, and Section 5.4 presents one particular reputation model, the Klein-Leffler model of product quality. Section 5.5 concludes the chapter with an overlapping generations model of consumer switching costs which uses the idea of Markov strategies to narrow down the number of equilibria.

The Chainstore Paradox

Suppose that we repeat Entry Deterrence I 20 times in the context of a chainstore that is trying to deter entry into 20 markets where it has outlets. We have seen that entry into just one market would not be deterred, but perhaps with 20 markets the outcome is different because the chainstore would fight the first entrant to deter the next 19.

The repeated game is much more complicated than the **one-shot game**, as the unrepeated version is called. A player's action is still to *Enter* or *Stay Out*, to *Fight* or *Collude*, but his strategy is a potentially very complicated rule telling him what action to choose depending on what actions both players took in each of the previous periods. Even the five-round repeated Prisoner's Dilemma has a strategy set for each player with over two billion strategies, and the number of strategy profiles is even greater (Sugden [1986], p. 108).

The obvious way to solve the game is from the beginning, where there is the least past history on which to condition a strategy, but that is not the easy way. We have to follow Kierkegaard, who said, "Life can only be understood backwards, but it must be lived forwards" (Kierkegaard 1938, p. 465). In picking his first action, a player looks ahead to its implications for all the future periods, so it is easiest to start by understanding the end of a multi- period game, where the future is shortest.

Consider the situation in which 19 markets have already been invaded (and maybe the chainstore fought, or maybe not). In the last market, the subgame in which the two players find themselves is identical to the one-shot Entry Deterrence I, so the entrant will *Enter* and the chainstore will *Collude*, regardless of the past history of the game. Next, consider the next- to-last market. The chainstore can gain nothing from building a reputation for ferocity, because it is common knowledge that he will *Collude* with the last entrant anyway. So he might as well *Collude* in the 19th market. But we can say the same of the 18th market and— by continuing backward induction— of every market, including the first. This result is called the **Chainstore Paradox** after Selten (1978) .

Backward induction ensures that the strategy profile is a subgame perfect equilibrium. There are other Nash equilibria— (*Always Fight, Never Enter*), for example— but because of the Chainstore Paradox they are not perfect.

The Repeated Prisoner's Dilemma

The Prisoner's Dilemma is similar to Entry Deterrence I. Here the prisoners would like to commit themselves to *Silence*, but, in the absence of commitment, they *Blame*. The Chainstore Paradox can be applied to show that repetition does not induce cooperative behavior. Both prisoners know that in the last repetition, both will *Blame*. After 18 repetitions, they know that no matter what happens in the 19th, both will *Blame* in the 20th, so they might as well *Blame* in the 19th too. Building a reputation is pointless, because in the 20th period it is not going to matter. Proceeding inductively, both players *Blame* in every period, the unique perfect equilibrium outcome.

In fact, as a consequence of the fact that the one-shot Prisoner's Dilemma has a dominant-strategy equilibrium, blaming is the only Nash outcome for the repeated Prisoner's Dilemma, not just the only perfect outcome. The argument of the previous paragraph did not show that blaming was the unique Nash outcome. To show subgame perfectness, we worked back from the end using longer and longer subgames. To show that blaming is the only Nash outcome, we do not look at subgames, but instead rule out successive classes of strategies from being Nash. Consider the portions of the strategy which apply to the equilibrium path (that is, the portions directly relevant to the payoffs). No strategy in the class that calls for *Silence* in the last period can be a Nash strategy, because the same strategy with *Blame* replacing *Silence* would dominate it. But if both players have strategies calling for blaming in the last period, then no strategy that does not call for blaming in the next-to-last period is Nash, because a player should deviate by replacing *Silence* with *Blame* in the next- to-last period. The argument can be carried back to the first period, ruling out any class of strategies that does not call for blaming everywhere along the equilibrium path.

The strategy of always blaming is not a dominant strategy, as it is in the one-shot game, because it is not the best response to various suboptimal strategies such as (*Silence until the other player Blamees, then Silence for the rest of the game*). Moreover, the uniqueness is only on the equilibrium path. Nonperfect Nash strategies could call for cooperation at nodes far away from the equilibrium path, since that action would never have to be taken. If Row has chosen (*Always Blame*), one of Column's best responses is (*Always Blame unless Row has chosen Silence ten times; then always Silence*).

5.2 Infinitely Repeated Games, Minimax Punishments, and the Folk Theorem

The contradiction between the Chainstore Paradox and what many people think of as real world behavior has been most successfully resolved by adding incomplete information to the model, as will be seen in Section 6.4. Before we turn to incomplete information, however, we will explore certain other modifications. One idea is to repeat the Prisoner's Dilemma an infinite number of times instead of a finite number (after all, few economies have a known end date). Without a last period, the inductive argument in the Chainstore Paradox fails.

In fact, we can find a simple perfect equilibrium for the infinitely repeated Prisoner's Dilemma in which both players cooperate—a game in which both players adopt the Grim Strategy.

Grim Strategy
1 Start by choosing Silence.
2 Continue to choose Silence unless some player has chosen Blame, in which case choose Blame forever.

Notice that the Grim Strategy says that even if a player is the first to deviate and choose *Blame*, he continues to choose *Blame* thereafter.

If Column uses the Grim Strategy, the Grim Strategy is weakly Row's best response. If Row cooperates, he will continue to receive the high (*Silence*, *Silence*) payoff forever. If he blamees, he will receive the higher (*Blame*, *Silence*) payoff once, but the best he can hope for thereafter is the (*Blame*, *Blame*) payoff.

Even in the infinitely repeated game, cooperation is not immediate, and not every strategy that punishes blaming is perfect. A notable example is the strategy of Tit-for-Tat.

Tit-for-Tat
1 Start by choosing Silence.
2 Thereafter, in period n choose the action that the other player chose in period $(n-1)$.

If Column uses Tit-for-Tat, Row does not have an incentive to *Blame* first, because if Row cooperates he will continue to receive the high (*Silence*, *Silence*) payoff, but if he blamees and then returns to Tit-for-Tat, the players alternate (*Blame*, *Silence*) with (*Silence*, *Blame*) forever. Row's average payoff from this alternation would be lower than if he had stuck to (*Silence*, *Silence*), and would swamp the one-time gain. But Tit-for-Tat is almost never perfect in the infinitely repeated Prisoner's Dilemma without discounting, because it is not rational for Column to punish Row's initial *Blame*. Adhering to Tit-for-Tat's punishments results in a miserable alternation of *Blame* and *Silence*, so Column would rather ignore Row's first *Blame*. The deviation is not from the equilibrium path action of *Silence*, but from the off-equilibrium action rule of *Blame in response to a Blame*. Thus, Tit-for-Tat, unlike the Grim Strategy, is not subgame perfect. (See Kalai, Samet & Stanford (1988) and Problem 5.5 for more on this point.)

Unfortunately, although eternal cooperation is a perfect equilibrium outcome in the infinite game under at least one strategy, so is practically anything else, including eternal blaming. The multiplicity of equilibria is summarized by the Folk Theorem, so called because its origins are hazy.[1]

Theorem 1 (the Folk Theorem)
In an infinitely repeated n-person game with finite action sets at each repetition, any profile of actions observed in any finite number of repetitions is the unique outcome of some subgame perfect equilibrium given
Condition 1: *The rate of time preference is zero, or positive and sufficiently small;*
Condition 2: *The probability that the game ends at any repetition is zero, or positive and sufficiently small; and*
Condition 3: *The set of payoff profiles that strictly Pareto dominate the minimax payoff profiles in the mixed extension of the one-shot game is n- dimensional.*

What the Folk Theorem tells us is that claiming that particular behavior arises in a perfect equilibrium is meaningless in an infinitely repeated game. This applies to any game that meets conditions 1 to 3, not just to the Prisoner's Dilemma. If an infinite amount of time always remains in the game, a way can always be found to make one player willing to punish some other player for the sake of a better future, even if the punishment currently hurts the punisher as well as the punished. Any finite interval of time is insignificant compared to eternity, so the threat of future reprisal makes the players willing to carry out the punishments needed.

We will next discuss conditions 1 to 3.

Condition 1: Discounting

The Folk Theorem helps answer the question of whether discounting future payments lessens the influence of the troublesome Last Period. Quite to the contrary, with discounting, the present gain from blaming is weighted more heavily and future gains from cooperation more lightly. If the discount rate is very high the game almost returns to being one-shot. When the real interest rate is 1,000 percent, a payment next year is little better than a payment a hundred years hence, so next year is practically irrelevant. Any model that relies on a large number of repetitions also assumes that the discount rate is not too high.

Allowing a little discounting is none the less important to show there is no discontinuity

[1]There is a multiplicity of Folk Theorems too, since the idea can be formalized in many ways and in many settings— infinitely or finitely repeated games, complete or incomplete information, overlapping generations, epsilon-equilibria, and so forth. Benoit & Krishna (2000) attempt a synthesis, using the basic principle that many things can happen if there is not an end-game that can pin down the equilibrium.

at the discount rate of zero. If we come across an undiscounted, infinitely repeated game with many equilibria, the Folk Theorem tells us that adding a low discount rate will not reduce the number of equilibria. This contrasts with the effect of changing the model by having a large but finite number of repetitions, a change which often eliminates all but one outcome by inducing the Chainstore Paradox.

A discount rate of zero supports many perfect equilibria, but if the rate is high enough, the only equilibrium outcome is eternal blaming. We can calculate the critical value for given parameters. The Grim Strategy imposes the heaviest possible punishment for deviant behavior. Using the payoffs for the Prisoner's Dilemma from Table 2a in the next section, the equilibrium payoff from the Grim Strategy is the current payoff of 5 plus the value of the rest of the game, which from Table 2 of Chapter 4 is $\frac{5}{r}$. If Row deviated by blaming, he would receive a current payoff of 10, but the value of the rest of the game would fall to 0. The critical value of the discount rate is found by solving the equation $5 + \frac{5}{r} = 10 + 0$, which yields $r = 1$, a discount rate of 100 percent or a discount factor of $\delta = 0.5$. Unless the players are extremely impatient, blaming is not much of a temptation.

Condition 2: A probability of the game ending

Time preference is fairly straightforward, but what is surprising is that assuming that the game ends in each period with probability θ does not make a drastic difference. In fact, we could even allow θ to vary over time, so long as it never became too large. If $\theta > 0$, the game ends in finite time with probability one; or, put less dramatically, the expected number of repetitions is finite, but it still behaves like a discounted infinite game, because the expected number of future repetitions is always large, no matter how many have already occurred. The game still has no Last Period, and it is still true that imposing one, no matter how far beyond the expected number of repetitions, would radically change the results.

The following two situations are different from each other.
"1 The game will end at some uncertain date before T."
"2 There is a constant probability of the game ending."

In situation (1), the game is like a finite game, because, as time passes, the maximum length of time still to run shrinks to zero. In situation (2), even if the game will end by T with high probability, if it actually lasts until T the game looks exactly the same as at time zero. The fourth verse from the hymn "Amazing grace" puts this stationarity very nicely (though I expect it is supposed to apply to a game with $\theta = 0$).

When we've been there ten thousand years,
Bright shining as the sun,
We've no less days to sing God's praise
Than when we'd first begun.

Condition 3: Dimensionality

The "minimax payoff" mentioned in Theorem 1 is the payoff that results if all the other players pick strategies solely to punish player i, and he protects himself as best he can.

The set of strategies s_{-i}^{i*} is a set of $(n-1)$ **minimax strategies** *chosen by all the players except i to keep i's payoff as low as possible, no matter how he responds. s_{-i}^{i*} solves*

$$\underset{s_{-i}}{\mathit{Minimize}}\ \underset{s_i}{\mathit{Maximum}}\ \pi_i(s_i, s_{-i}). \tag{1}$$

Player i's **minimax payoff**, **minimax value**, *or* **security value** *is his payoff from the solution of (1).*

The dimensionality condition is needed only for games with three or more players. It is satisfied if there is some payoff profile for each player in which his payoff is greater than his minimax payoff but still different from the payoff of every other player. Figure 1 shows how this condition is satisfied for the two-person Prisoner's Dilemma of Table 2a a few pages beyond this paragraph, but not for the two-person Ranked Coordination game. It is also satisfied by the n-person Prisoner's Dilemma in which a solitary blameer gets a higher payoff than his cooperating fellow-prisoners, but not by the n- person Ranked Coordination game, in which all the players have the same payoff. The condition is necessary because establishing the desired behavior requires some way for the other players to punish a deviator without punishing themselves.

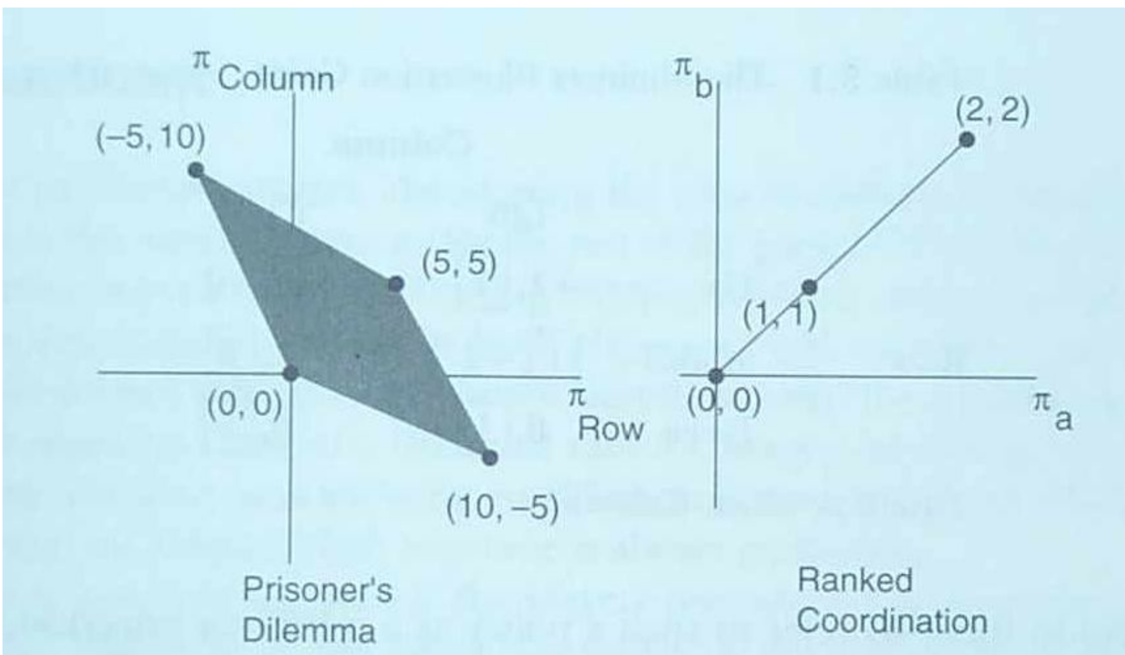

Figure 1: The Dimensionality Condition

An alternative to the dimensionality condition in the Folk Theorem is

Condition 3′: *The repeated game has a "desirable" subgame-perfect equilibrium in which the strategy profile $\overline{s}$ played each period gives player i a payoff that exceeds his payoff from some other "punishment" subgame-perfect equilibrium in which the strategy profile $\underline{s}^i$ is played each period:*

$$\exists \overline{s} : \forall i,\ \exists \underline{s}^i : \pi_i(\underline{s}^i) < \pi_i(\overline{s}).$$

Condition 3′ is useful because sometimes it is easy to find a few perfect equilibria. To enforce the desired pattern of behavior, use the "desirable" equilibrium as a carrot and the "punishment" equilibrium as a self-enforcing stick (see Rasmusen [1992a]).

Minimax and Maximin

In discussions of strategies which enforce cooperation, the question of the maximum severity of punishment strategies frequently arises. Thus, the idea of the minimax strategy—the most severe sanction possible if the offender does not cooperate in his own punishment— entered into the statement of the Folk Theorem. The corresponding strategy for an offender trying to protect himself from punishment, is the maximin strategy.

The strategy s_i^* is a **maximin strategy** *for player i if, given that the other players pick strategies to make i's payoff as low as possible, s_i^* gives i the highest possible payoff. In our notation, s_i^* solves*

$$\underset{s_i}{Maximize} \;\; \underset{s_{-i}}{Minimum} \;\; \pi_i(s_i, s_{-i}). \tag{2}$$

The following formulas show how to calculate the minimax and maximin strategies for a two-player game with Player 1 as i.

$$\text{Maximin:} \quad \underset{s_1}{Maximum} \quad \underset{s_2}{Minimum} \quad \pi_1$$

$$\text{Minimax:} \quad \underset{s_2}{Minimum} \quad \underset{s_1}{Maximum} \quad \pi_1$$

In the Prisoner's Dilemma, the minimax and maximin strategies are both *Blame*. Although the Welfare Game (Chapter 3's Table 1) has only a mixed strategy Nash equilibrium, if we restrict ourselves to the pure strategies (just for illustration here) the Pauper's maximin strategy is *Try to Work*, which guarantees him at least 1, and his strategy for minimaxing the Government is *Be Idle*, which prevents the Government from getting more than zero.

Under minimax, Player 2 is purely malicious but must move first (at least in choosing a mixing probability) in his attempt to cause player 1 the maximum pain. Under maximin, Player 1 moves first, in the belief that Player 2 is out to get him. In variable-sum games, minimax is for sadists and maximin for paranoids. In zero-sum games, the players are merely neurotic. Minimax is for optimists, and maximin is for pessimists.

The maximin strategy need not be unique, and it can be in mixed strategies. Since maximin behavior can also be viewed as minimizing the maximum loss that might be suffered, decision theorists refer to such a policy as a **minimax criterion,** a catchier phrase (Luce & Raiffa [1957], p. 279).

It is tempting to use maximin strategies as the basis of an equilibrium concept. A **maximin equilibrium** is made up of a maximin strategy for each player. Such a strategy might seem reasonable because each player then has protected himself from the worst harm possible. Maximin strategies have very little justification, however, for a rational player. They are not simply the optimal strategies for risk-averse players, because risk aversion is accounted for in the utility payoffs. The players' implicit beliefs can be inconsistent in a maximin equilibrium, and a player must believe that his opponent would choose the most harmful strategy out of spite rather than self-interest if maximin behavior is to be rational.

The usefulness of minimax and maximin strategies is not in directly predicting the best strategies of the players, but in setting the bounds of how their strategies affect their payoffs, as in condition 3 of Theorem 1.

It is important to remember that minimax and maximin strategies are not always pure strategies. In the Minimax Illustration Game of Table 1, which I take from Fudenerg & Tirole (1991a, p. 150), Row can guarantee himself a payoff of 0 by choosing *Down*, so that is his maximin strategy. Column cannot hold Row's payoff down to 0, however, by using a pure minimax strategy. If Column chooses *Left*, Row can choose *Middle* and get a payoff of 1; if Column chooses *Right*, Row can choose *Up* and get a payoff of 1. Column can, however, hold Row's payoff down to 0 by choosing a mixed minimax strategy of *(Probability 0.5 of Left, Probability 0.5 of Right)*. Row would then respond with *Down*, for a minimax payoff of 0, since either *Up*, *Middle*, or a mixture of the two would give him a payoff of -0.5 $(= 0.5(-2) + 0.5(1))$.

Table 1: The Minimax Illustration Game

		Column	
		Left	*Right*
	Up	$-2, \boxed{2}$	$\boxed{1}, -2$
Row:	*Middle*	$\boxed{1}, -2$	$-2, \boxed{2}$
	Down	$0, \boxed{1}$	$0, \boxed{1}$

Payoffs to: (Row, Column). Best-response payoffs are boxed.

Column's maximin and minimax strategies can also be computed. The strategy for minimaxing Column is *(Probability 0.5 of Up, Probability 0.5 of Middle)*, his maximin strategy is *(Probability 0.5 of Left, Probability 0.5 of Right)*, and his minimax payoff is 0.

In two-person zero-sum games, minimax and maximin strategies are more directly useful, because when Player 1 reduces Player 2's payoff, he increases his own payoff. Punishing the other player is equivalent to rewarding yourself. This is the origin of the celebrated **Minimax Theorem** (von Neumann [1928]), which says that a minimax equilibrium exists in pure or mixed strategies for every two-person zero-sum game and is identical to the maximin equilibrium. Unfortunately, the games that come up in applications are usually not zero-sum games, so the Minimax Theorem usually cannot be applied.

Precommitment

What if we use metastrategies, abandoning the idea of perfectness by allowing players to commit at the start to a strategy for the rest of the game? We would still want to keep the game noncooperative by disallowing binding promises, but we could model it as a game with simultaneous choices by both players, or with one move each in sequence.

If precommitted strategies are chosen simultaneously, the equilibrium outcome of the finitely repeated Prisoner's Dilemma calls for always blaming, because allowing commitment is the same as allowing equilibria to be nonperfect, in which case, as was shown earlier, the unique Nash outcome is always blaming.

A different result is achieved if the players precommit to strategies in sequence. The outcome depends on the particular values of the parameters, but one possible equilibrium is the following: Row moves first and chooses the strategy (*Silence* until Column *Blame*s; thereafter always *Blame*), and Column chooses (*Silence* until the last period; then *Blame*). The observed outcome would be for both players to choose *Silence* until the last period, and then for Row to again choose *Silence* , but for Column to choose *Blame* . Row would submit to this because if he chose a strategy that initiated blaming earlier, Column would choose a strategy of starting to blame earlier too. The game has a second-mover advantage.

5.3 Reputation: The One-Sided Prisoner's Dilemma

Part II of this book will analyze moral hazard and adverse selection. Under moral hazard, a player wants to commit to high effort, but he cannot credibly do so. Under adverse selection, a player wants to prove he is high ability, but he cannot. In both, the problem is that the penalties for lying are insufficient. Reputation seems to offer a way out of the problem. If the relationship is repeated, perhaps a player is willing to be honest in early periods in order to establish a reputation for honesty which will be valuable to himself later.

Reputation seems to play a similar role in making threats to punish credible. Usually punishment is costly to the punisher as well as the punished, and it is not clear why the punisher should not let bygones be bygones. Yet in 1988 the Soviet Union paid off 70-year-old debt to dissuade the Swiss authorities from blocking a mutually beneficial new bond issue ("Soviets Agree to Pay Off Czarist Debt to Switzerland," *Wall Street Journal*, January 19, 1988, p. 60). Why were the Swiss so vindictive towards Lenin?

The questions of why players do punish and do not cheat are really the same questions that arise in the repeated Prisoner's Dilemma, where the fact of an infinite number of repetitions allows cooperation. That is the great problem of reputation. Since everyone knows that a player will *Blame*, choose low effort, or default on debt in the last period, why do they suppose he will bother to build up a reputation in the present? Why should past behavior be any guide to future behavior?

Not all reputation problems are quite the same as the Prisoner's Dilemma, but they have much the same flavor. Some games, like duopoly or the original Prisoner's Dilemma, are **two-sided** in the sense that each player has the same strategy set and the payoffs are symmetric. Others, such as the game of Product Quality (see below), are what we might call **one-sided Prisoner's Dilemmas**, which have properties similar to the Prisoner's Dilemma, but do not fit the usual definition because they are asymmetric. Table 2 shows the normal forms for both the original Prisoner's Dilemma and the one- sided version.[2] The important difference is that in the one-sided Prisoner's Dilemma at least one player

[2]The exact numbers are different from the Prisoner's Dilemma in Table 1 in Chapter 1, but the ordinal rankings are the same. Numbers such as those in Table 2 of the present chapter are more commonly used,

really does prefer the outcome equivalent to (*Silence*, *Silence*), which is (*High Quality, Buy)* in Table 2b, to anything else. He blames defensively, rather than both offensively and defensively. The payoff (0,0) can often be interpreted as the refusal of one player to interact with the other, for example, the motorist who refuses to buy cars from Chrysler because he knows they once falsified odometers. Table 3 lists examples of both one-sided and two-sided games.

Table 2: Prisoner's Dilemmas

(a) Two-Sided (conventional)

		Column		
		Silence		*Blame*
	Silence	5,5	→	-5,10
Row:		↓		↓
	Blame	10,-5	→	**0,0**

Payoffs to: (Row, Column). Arrows show how a player can increase his payoff.

(b) One-Sided

		Consumer (Column)		
		Buy		*Boycott*
	High Quality	5,5	←	0,0
Seller (Row):		↓		↕
	Low Quality	10, -5	→	**0,0**

Payoffs to: (Seller, Consumer). Arrows show how a player can increase his payoff.

Table 3 Some Repeated Games in which Reputation Is Important

because it is convenient to normalize the (*Blame, Blame*) payoffs to (0,0) and to make most of the numbers positive rather than negative.

Application	Sidedness	Players	Actions
Prisoner's Dilemma	two-sided	Row	*Silence/Blame*
		Column	*Silence/Blame*
Duopoly	two-sided	Firm	*High price/Low price*
		Firm	*High price/Low price*
Employment	two-sided	Employer	*Bonus/No bonus*
		Employee	*Work/Shirk*
Product Quality	one-sided	Consumer	*Buy/Boycott*
		Seller	*High quality/low quality*
Entry Deterrence	one-sided	Incumbent	*Low price/High price*
		Entrant	*Enter/Stay out*
Financial Disclosure	one-sided	Corporation	*Truth/Lies*
		Investor	*Invest/Refrain*
Borrowing	one-sided	Lender	*Lend/Refuse*
		Borrower	*Repay/Default*

The Nash and iterated dominance equilibria in the one-sided Prisoner's Dilemma are still (*Blame, Blame*), but it is not a dominant-strategy equilibrium. Column does not have a dominant strategy, because if Row were to choose *Silence*, Column would also choose *Silence*, to obtain the payoff of 5; but if Row chooses *Blame*, Column would choose *Blame*, for a payoff of zero. *Blame* is however, weakly dominant for Row, which makes $(Blame, Blame)$ the iterated dominant strategy equilibrium. In both games, the players would like to persuade each other that they will cooperate, and devices that induce cooperation in the one-sided game will usually obtain the same result in the two-sided game.

5.4 Product Quality in an Infinitely Repeated Game

The Folk Theorem tells us that some perfect equilibrium of an infinitely repeated game— sometimes called an **infinite horizon model**— can generate any pattern of behavior observed over a finite number of periods. But since the Folk Theorem is no more than a mathematical result, the strategies that generate particular patterns of behavior may be unreasonable. The theorem's value is in provoking close scrutiny of infinite horizon models so that the modeller must show why his equilibrium is better than a host of others. He must go beyond satisfaction of the technical criterion of perfectness and justify the strategies on other grounds.

In the simplest model of product quality, a seller can choose between producing costly high quality or costless low quality, and the buyer cannot determine quality before he purchases. If the seller would produce high quality under symmetric information, we have a one-sided Prisoner's Dilemma, as in Table 2b. Both players are better off when the seller produces high quality and the buyer purchases the product, but the seller's weakly dominant strategy is to produce low quality, so the buyer will not purchase. This is also an example of moral hazard, the topic of chapter 7.

A potential solution is to repeat the game, allowing the firm to choose quality at each repetition. If the number of repetitions is finite, however, the outcome stays the same because of the Chainstore Paradox. In the last repetition, the subgame is identical to the one- shot game, so the firm chooses low quality. In the next-to-last repetition, it is foreseen that the last period's outcome is independent of current actions, so the firm also chooses low quality, an argument that can be carried back to the first repetition.

If the game is repeated an infinite number of times, the Chainstore Paradox is inapplicable and the Folk Theorem says that a wide range of outcomes can be observed in equilibrium. Klein & Leffler (1981) construct a plausible equilibrium for an infinite period model. Their original article, in the traditional verbal style of UCLA, does not phrase the result in terms of game theory, but we will recast it here, as I did in Rasmusen (1989b). In equilibrium, the firm is willing to produce a high quality product because it can sell at a high price for many periods, but consumers refuse to ever buy again from a firm that has once produced low quality. The equilibrium price is high enough that the firm is unwilling to sacrifice its future profits for a one- time windfall from deceitfully producing low quality and selling it at a high price. Although this is only one of a large number of subgame perfect equilibria, the consumers' behavior is simple and rational: no consumer can benefit by deviating from the equilibrium.

Product Quality

Players
An infinite number of potential firms and a continuum of consumers.

The Order of Play
1 An endogenous number n of firms decide to enter the market at cost F.
2 A firm that has entered chooses its quality to be *High* or *Low*, incurring the constant marginal cost c if it picks *High* and zero if it picks *Low*. The choice is unobserved by consumers. The firm also picks a price p.
3 Consumers decide which firms (if any) to buy from, choosing firms randomly if they are indifferent. The amount bought from firm i is denoted q_i.
4 All consumers observe the quality of all goods purchased in that period.
5 The game returns to (2) and repeats.

Payoffs
The consumer benefit from a product of low quality is zero, but consumers are willing to buy quantity $q(p) = \sum_{i=1}^n q_i$ for a product believed to be high quality, where $\frac{dq}{dp} < 0$.
If a firm stays out of the market, its payoff is zero.
If firm i enters, it receives $-F$ immediately. Its current end-of- period payoff is $q_i p$ if it produces *Low* quality and $q_i(p - c)$ if it produces *High* quality. The discount rate is $r \geq 0$.

That the firm can produce low quality items at zero marginal cost is unrealistic, but it is only a simplifying assumption. By normalizing the cost of producing low quality to zero, we avoid having to carry an extra variable through the analysis without affecting the result.

The Folk Theorem tells us that this game has a wide range of perfect outcomes, including a large number with erratic quality patterns like (*High, High, Low, High, Low, Low*...). If we confine ourselves to pure- strategy equilibria with the stationary outcome of constant quality and identical behavior by all firms in the market, then the two outcomes are low quality and high quality. Low quality is always an equilibrium outcome, since it is an equilibrium of the one-shot game. If the discount rate is low enough, high quality is also an equilibrium outcome, and this will be the focus of our attention. Consider the following strategy profile:

Firms. $\tilde{n}$ firms enter. Each produces high quality and sells at price $\tilde{p}$. If a firm ever deviates from this, it thereafter produces low quality (and sells at the same price $\tilde{p}$). The values of $\tilde{p}$ and $\tilde{n}$ are given by equations (4) and (8) below.

Buyers. Buyers start by choosing randomly among the firms charging $\tilde{p}$. Thereafter, they remain with their initial firm unless it changes its price or quality, in which case they switch randomly to a firm that has not changed its price or quality.

This strategy profile is a perfect equilibrium. Each firm is willing to produce high quality and refrain from price-cutting because otherwise it would lose all its customers. If it has deviated, it is willing to produce low quality because the quality is unimportant, given the absence of customers. Buyers stay away from a firm that has produced low quality because they know it will continue to do so, and they stay away from a firm that has cut the price because they know it will produce low quality. For this story to work, however, the equilibrium must satisfy three constraints that will be explained in more depth in Section 7.3: incentive compatibility, competition, and market clearing.

The **incentive compatibility** constraint says that the individual firm must be willing to produce high quality. Given the buyers' strategy, if the firm ever produces low quality it receives a one-time windfall profit, but loses its future profits. The tradeoff is represented by constraint (3), which is satisfied if the discount rate is low enough.

$$\frac{q_i p}{1+r} \leq \frac{q_i(p-c)}{r} \qquad (\textit{incentive compatibility}). \tag{3}$$

Inequality (3) determines a lower bound for the price, which must satisfy

$$\tilde{p} \geq (1+r)c. \tag{4}$$

Condition (4) will be satisfied as an equality, because any firm trying to charge a price higher than the quality- guaranteeing $\tilde{p}$ would lose all its customers.

The second constraint is that competition drives profits to zero, so firms are indifferent between entering and staying out of the market.

$$\frac{q_i(p-c)}{r} = F \qquad (\textit{competition}) \tag{5}$$

Treating (3) as an equation and using it to replace p in equation (5) gives

$$q_i = \frac{F}{c}. \tag{6}$$

We have now determined p and q_i, and only n remains, which is determined by the equality of supply and demand. The market does not always clear in models of asymmetric information (see Stiglitz [1987]), and in this model each firm would like to sell more than its equilibrium output at the equilibrium price, but the market output must equal the quantity demanded by the market.

$$nq_i = q(p). \qquad (market\ \ clearing) \tag{7}$$

Combining equations (3), (6), and (7) yields

$$\tilde{n} = \frac{cq([1+r]c)}{F}. \tag{8}$$

We have now determined the equilibrium values, the only difficulty being the standard existence problem caused by the requirement that the number of firms be an integer (see note N5.4).

The equilibrium price is fixed because F is exogenous and demand is not perfectly inelastic, which pins down the size of firms. If there were no entry cost, but demand were still elastic, then the equilibrium price would still be the unique p that satisfied constraint (3), and the market quantity would be determined by $q(p)$, but F and q_i would be undetermined. If consumers believed that any firm which might possibly produce high quality paid an exogenous dissipation cost F, the result would be a continuum of equilibria. The firms' best response would be for $\tilde{n}$ of them to pay F and produce high quality at price $\tilde{p}$, where $\tilde{n}$ is determined by the zero profit condition as a function of F. Klein & Leffler note this indeterminacy and suggest that the profits might be dissipated by some sort of brand-specific capital. This is especially plausible when there is asymmetric information, so firms might wish to use capital spending to signal that they intend to be in the business for a long time; Rasmusen & Perri (2001) shows a way to model this. Another good explanation for which firms enjoy the high profits of good reputation is simply the history of the industry. Schmalensee (1982) shows how a pioneering brand can retain a large market share because consumers are unwilling to investigate the quality of new brands.

The repeated-game model of reputation for product quality can be used to model many other kinds of reputation too. Even before Klein & Leffler (1981), Telser titled his 1980 article "A Theory of Self- Enforcing Agreements," and looked at a number of situations in which repeated play balanced the short-run gain from cheating against the long-run gain from cooperation. We will see the idea later in this book in in Section 8.1 as part of the idea of the "efficiency wage".

Keep in mind, however, that "reputation" can be modelled in two distinct ways. In our next model, a firm with a good reputation is one which produces high quality to avoid losing that reputation, a "moral hazard" model because the focus is on the player's choice of actions. An alternative is a model in which the firm with a good reputation is one which has shown that it would not produce low quality even if there were no adverse consequences from doing so, an "adverse selection" model because the focus is on the player's type. One kind of reputation is for deciding to be good; the other is for Nature having chosen the player to be good. As you will see, the Gang of Four model of Chapter 6 will mix the two.

*5.5 Markov Equilibria and Overlapping Generations: Customer Switching Costs

The next model demonstrates a general modelling technique, the **overlapping generations model**, in which different cohorts of otherwise identical players enter and leave the game with overlapping "lifetimes," and a new equilibrium concept, "Markov equilibrium." The best-known example of an overlapping-generations model is the original consumption-loans model of Samuelson (1958). The models are most often used in macroeconomics, but they can also be useful in microeconomics. Klemperer (1987) has stimulated considerable interest in customers who incur costs in moving from one seller to another. The model used here will be that of Farrell & Shapiro (1988).

Customer Switching Costs

Players
Firms Apex and Brydox, and a series of customers, each of whom is first called a youngster and then an oldster.

The Order of Play
1a Brydox, the initial incumbent, picks the incumbent price p_1^i.
1b Apex, the initial entrant, picks the entrant price p_1^e.
1c The oldster picks a firm.
1d The youngster picks a firm.
1e Whichever firm attracted the youngster becomes the incumbent.
1f The oldster dies and the youngster becomes an oldster.
2a Return to (1a), possibly with new identities for entrant and incumbent.

Payoffs
The discount factor is δ. The customer reservation price is R and the switching cost is c. The per period payoffs in period t are, for $j = (i, e)$,

$$\pi_{firm\ j} = \begin{cases} 0 & \text{if no customers are attracted.} \\ p_t^j & \text{if just oldsters or just youngsters are attracted.} \\ 2p_t^j & \text{if both oldsters and youngsters are attracted.} \end{cases}$$

$$\pi_{oldster} = \begin{cases} R - p_t^i & \text{if he buys from the incumbent.} \\ R - p_t^e - c & \text{if he switches to the entrant.} \end{cases}$$

$$\pi_{youngster} = \begin{cases} R - p_t^i & \text{if he buys from the incumbent.} \\ R - p_t^e & \text{if he buys from the entrant.} \end{cases}$$

Finding all the perfect equilibria of an infinite game like this one is difficult, so we will follow Farrell and Shapiro in limiting ourselves to the much easier task of finding the perfect Markov equilibrium, which is unique.

A **Markov strategy** *is a strategy that, at each node, chooses the action independently of the history of the game except for the immediately preceding action (or actions, if they were simultaneous).*

Here, a firm's Markov strategy is its price as a function of whether the particular is the incumbent or the entrant, and not a function of the entire past history of the game.

There are two ways to use Markov strategies: (1) just look for equilibria that use Markov strategies, and (2) disallow nonMarkov strategies and then look for equilibria. Because the first way does not disallow non-Markov strategies, the equilibrium must be such that no player wants to deviate by using any other strategy, whether Markov or not. This is just a way of eliminating possible multiple equilibria by discarding ones that use non-Markov strategies. The second way is much more dubious, because it requires the players not to use non- Markov strategies, even if they are best responses. A **perfect Markov equilibrium** uses the first approach: it is a perfect equilibrium that happens to use only Markov strategies.

Brydox, the initial incumbent, moves first and chooses p^i low enough that Apex is not tempted to choose $p^e < p^i - c$ and steal away the oldsters. Apex's profit is p^i if it chooses $p^e = p^i$ and serves just youngsters, and $2(p^i - c)$ if it chooses $p^e = p^i - c$ and serves both oldsters and youngsters. Brydox chooses p^i to make Apex indifferent between these alternatives, so

$$p^i = 2(p^i - c), \tag{9}$$

and

$$p^i = p^e = 2c. \tag{10}$$

In equilibrium, Apex and Brydox take turns being the incumbent and charge the same price.

Because the game lasts forever and the equilibrium strategies are Markov, we can use a trick from dynamic programming to calculate the payoffs from being the entrant versus being the incumbent. The equilibrium payoff of the current entrant is the immediate payment of p^e plus the discounted value of being the incumbent in the next period:

$$\pi_e^* = p^e + \delta\pi_i^*. \tag{11}$$

The incumbent's payoff can be similarly stated as the immediate payment of p^i plus the discounted value of being the entrant next period:

$$\pi_i^* = p^i + \delta\pi_e^*. \tag{12}$$

We could use equation (10) to substitute for p^e and p^i, which would leave us with the two equations (11) and (12) for the two unknowns π_i^* and π_e^*, but an easier way to compute the payoff is to realize that in equilibrium the incumbent and the entrant sell the same amount at the same price, so $\pi_i^* = \pi_e^*$ and equation (12) becomes

$$\pi_i^* = 2c + \delta\pi_i^*. \tag{13}$$

It follows that

$$\pi_i^* = \pi_e^* = \frac{2c}{1-\delta}. \tag{14}$$

Prices and total payoffs are increasing in the switching cost c, because that is what gives the incumbent market power and prevents ordinary competition of the ordinary Bertrand kind to be analyzed in section 13.2. The total payoffs are increasing in δ for the usual reason that future payments increase in value as δ approaches one.

*5.6 Evolutionary Equilibrium: Hawk-Dove

For most of this book we have been using the Nash equilibrium concept or refinements of it based on information and sequentiality, but in biology such concepts are often inappropriate. The lower animals are less likely than humans to think about the strategies of their opponents at each stage of a game. Their strategies are more likely to be preprogrammed and their strategy sets more restricted than the businessman's, if perhaps not more so than his customer's. In addition, behavior evolves, and any equilibrium must take account of the possibility of odd behavior caused by the occasional mutation. That the equilibrium is common knowledge, or that players cannot precommit to strategies, are not compelling assumptions. Thus, the ideas of Nash equilibrium and sequential rationality are much less useful than when game theory is modelling rational players.

Game theory has grown to some importance in biology, but the style is different than in economics. The goal is not to explain how players would rationally pick actions in a given situation, but to explain how behavior evolves or persists over time under exogenous shocks. Both approaches end up defining equilibria to be strategy profiles that are best responses in some sense, but biologists care much more about the stability of the equilibrium and how strategies interact over time. In section 3.5, we touched briefly on the stability of the Cournot equilibrium, but economists view stability as a pleasing by-product of the equilibrium rather than its justification. For biologists, stability is the point of the analysis.

Consider a game with identical players who engage in pairwise contests. In this special context, it is useful to think of an equilibrium as a strategy profile such that no player with a new strategy can enter the environment (**invade**) and receive a higher expected payoff than the old players. Moreover, the invading strategy should continue to do well even if it plays itself with finite probability, or its invasion could never grow to significance. In the commonest model in biology, all the players adopt the same strategy in equilibrium, called an evolutionarily stable strategy. John Maynard Smith originated this idea, which is somewhat confusing because it really aims at an equilibrium concept, which involves a strategy profile, not just one player's strategy. For games with pairwise interactions and identical players, however, the evolutionarily stable strategy can be used to define an equilibrium concept.

A strategy s^* is an **evolutionarily stable strategy***, or* **ESS***, if, using the notation $\pi(s_i, s_{-i})$ for player i's payoff when his opponent uses strategy s_{-i}, for every other strategy s' either*

$$\pi(s^*, s^*) > \pi(s', s^*) \tag{15}$$

or

$$\begin{array}{l} (a)\ \ \pi(s^*, s^*) = \pi(s', s^*) \\ \text{and} \\ (b)\ \ \pi(s^*, s') > \pi(s', s'). \end{array} \tag{16}$$

If condition (15) holds, then a population of players using s^* cannot be invaded by a deviant using s'. If condition (16) holds, then s' does well against s^*, but badly against itself, so that if more than one player tried to use s' to invade a population using s^*, the invaders would fail.

We can interpret ESS in terms of Nash equilibrium. Condition (15) says that s^* is a strong Nash equilibrium (although not every strong Nash strategy is an ESS). Condition (16) says that if s^* is only a weak Nash strategy, the weak alternative s' is not a best response to itself. ESS is a refinement of Nash, narrowed by the requirement that ESS not only be a best response, but that (a) it have the highest payoff of any strategy used in equilibrium (which rules out equilibria with asymmetric payoffs), and (b) it be a strictly best response to itself.

The motivations behind the two equilibrium concepts are quite different, but the similarities are useful because even if the modeller prefers ESS to Nash, he can start with the Nash strategies in his efforts to find an ESS.

As an example of (a), consider the Battle of the Sexes. In it, the mixed strategy equilibrium is an ESS, because a player using it has as high a payoff as any other player. The two pure strategy equilibria are not made up of ESS's, though, because in each of them one player's payoff is higher than the other's. Compare with Ranked Coordination, in which the two pure strategy equilibria and the mixed strategy equilibrium are all made up of ESS's. (The dominated equilibrium strategy is nonetheless an ESS, because given that the other players are using it, no player could do as well by deviating.)

As an example of (b), consider the Utopian Exchange Economy game in Table 4, adapted from problem 7.5 of Gintis (2000). In Utopia, each citizen can produce either one or two units of individualized output. He will then go into the marketplace and meet another citizen. If either of them produced only one unit, trade cannot increase their payoffs. If both of them produced two, however, they can trade one unit for one unit, and both end up happier with their increased variety of consumption.

Table 4 The Utopian Exchange Economy Game

			Jones	
		Low Output		*HighOutput*
	LowOutput	**1, 1**	$\leftrightarrow$	1, 1
Smith:		$\updownarrow$		$\downarrow$
	High Output	1,1	$\rightarrow$	**2,2**

Payoffs to: (Smith, Jones). Arrows show how a player can increase his payoff.

This game has three Nash equilibria, one of which is in mixed strategies. Since all strategies but *High Output* are weakly dominated, that alone is an ESS. *Low Output* fails to meet condition (16b), because it is not the strictly best response to itself. If the economy began with all citizens choosing *Low Output*, then if Smith deviated to *High Output* he would not do any better, but if *two* people deviated to *High Output*, they would do better in expectation because they might meet each other and receive the payoff of (2,2).

An Example of ESS: Hawk-Dove

The best-known illustration of the ESS is the game of Hawk-Dove . Imagine that we have a population of birds, each of whom can behave as an aggressive Hawk or a pacific Dove. We will focus on two randomly chosen birds, Bird One and Bird Two. Each bird has a choice of what behavior to choose on meeting another bird. A resource worth $V = 2$ "fitness units" is at stake when the two birds meet. If they both fight, the loser incurs a cost of $C = 4$, which means that the expected payoff when two Hawks meet is -1 $(= 0.5[2] + 0.5[-4])$ for each of them. When two Doves meet, they split the resource, for a payoff of 1 apiece. When a Hawk meets a Dove, the Dove flees for a payoff of 0, leaving the Hawk with a payoff of 2. Table 5 summarizes this.

Table 5 Hawk-Dove: Economics Notation

		Bird Two		
		Hawk		*Dove*
	Hawk	-1,-1	$\rightarrow$	**2,0**
Bird One:		$\downarrow$		$\uparrow$
	Dove	**0, 2**	$\leftarrow$	1,1

Payoffs to: (Bird One, Bird Two). Arrows show how a player can increase his payoff.

These payoffs are often depicted differently in biology games. Since the two players are identical, one can depict the payoffs by using a table showing the payoffs only of the row player. Applying this to Hawk-Dove generates Table 6.

Table 6 Hawk-Dove: Biology Notation

		Bird Two	
		Hawk	*Dove*
	Hawk	-1	2
Bird One:			
	Dove	0	1

Payoffs to: (Bird One)

Hawk-Dove is Chicken with new feathers. The two games have the same ordinal ranking of payoffs, as can be seen by comparing Table 5 with Chapter 3's Table 2, and their equilibria are the same except for the mixing parameters. Hawk-Dove has no symmetric pure-strategy Nash equilibrium, and hence no pure-strategy ESS, since in the two asymmetric Nash equilibria, *Hawk* gives a bigger payoff than *Dove*, and the doves would disappear from the population. In the ESS for this game, neither hawks nor doves completely take over the environment. If the population consisted entirely of hawks, a dove could invade and obtain a one-round payoff of 0 against a hawk, compared to the -1 that a hawk obtains against itself. If the population consisted entirely of doves, a hawk could invade and obtain a one-round payoff of 2 against a dove, compared to the 1 that a dove obtains against a dove.

In the mixed-strategy ESS, the equilibrium strategy is to be a hawk with probability 0.5 and a dove with probability 0.5, which can be interpreted as a population 50 percent hawks and 50 percent doves. As in the mixed-strategy equilibria in chapter 3, the players are indifferent as to their strategies. The expected payoff from being a hawk is the 0.5(2) from meeting a dove plus the 0.5(−1) from meeting another hawk, a sum of 0.5. The expected payoff from being a dove is the 0.5(1) from meeting another dove plus the 0.5(0) from meeting a hawk, also a sum of 0.5. Moreover, the equilibrium is stable in a sense similar to the Cournot equilibrium. If 60 percent of the population were hawks, a bird would have a higher fitness level as a dove. If "higher fitness" means being able to reproduce faster, the number of doves increases and the proportion returns to 50 percent over time.

The ESS depends on the strategy sets allowed the players. If two birds can base their behavior on commonly observed random events such as which bird arrives at the resource first, and $V < C$ (as specified above), then a strategy called the **bourgeois strategy** is an ESS. Under this strategy, the bird respects property rights like a good bourgeois; it behaves as a hawk if it arrives first, and a dove if it arrives second, where we assume the order of arrival is random. The bourgeois strategy has an expected payoff of 1 from meeting itself, and behaves exactly like a 50:50 randomizer when it meets a strategy that ignores the order of arrival, so it can successfully invade a population of 50:50 randomizers. But the bourgeois strategy is a correlated strategy (see section 3.3), and requires something like the order of arrival to decide which of two identical players will play *Hawk*.

The ESS is suited to games in which all the players are identical and interacting in pairs. It does not apply to games with non-identical players— wolves who can be wily or big and deer who can be fast or strong— although other equilibrium concepts of the same flavor can be constructed. The approach follows three steps, specifying (1) the initial population proportions and the probabilities of interactions, (2) the pairwise interactions, and (3) the dynamics by which players with higher payoffs increase in number in the population. Economics games generally use only the second step, which describes the strategies and payoffs from a single interaction.

The third step, the evolutionary dynamics, is especially foreign to economics. In specifying dynamics, the modeller must specify a difference equation (for discrete time) or differential equation (for continuous time) that describes how the strategies employed change over iterations, whether because players differ in the number of their descendants or because they learn to change their strategies over time. In economics games, the adjustment process is usually degenerate: the players jump instantly to the equilibrium. In biology games, the adjustment process is slower and cannot be derived from theory. How quickly the population of hawks increases relative to doves depends on the metabolism of the bird and the length of a generation.

Slow dynamics also makes the starting point of the game important, unlike the case when adjustment is instantaneous. Figure 2, taken from David Friedman (1991), shows a way to graphically depict evolution in a game in which all three strategies of *Hawk*, *Dove*, and *Bourgeois* are used. A point in the triangle represents a proportion of the three strategies in the population. At point E_3, for example, half the birds play *Hawk*, half play *Dove*, and none play *Bourgeois*, while at E_4 all the birds play *Bourgeois*.

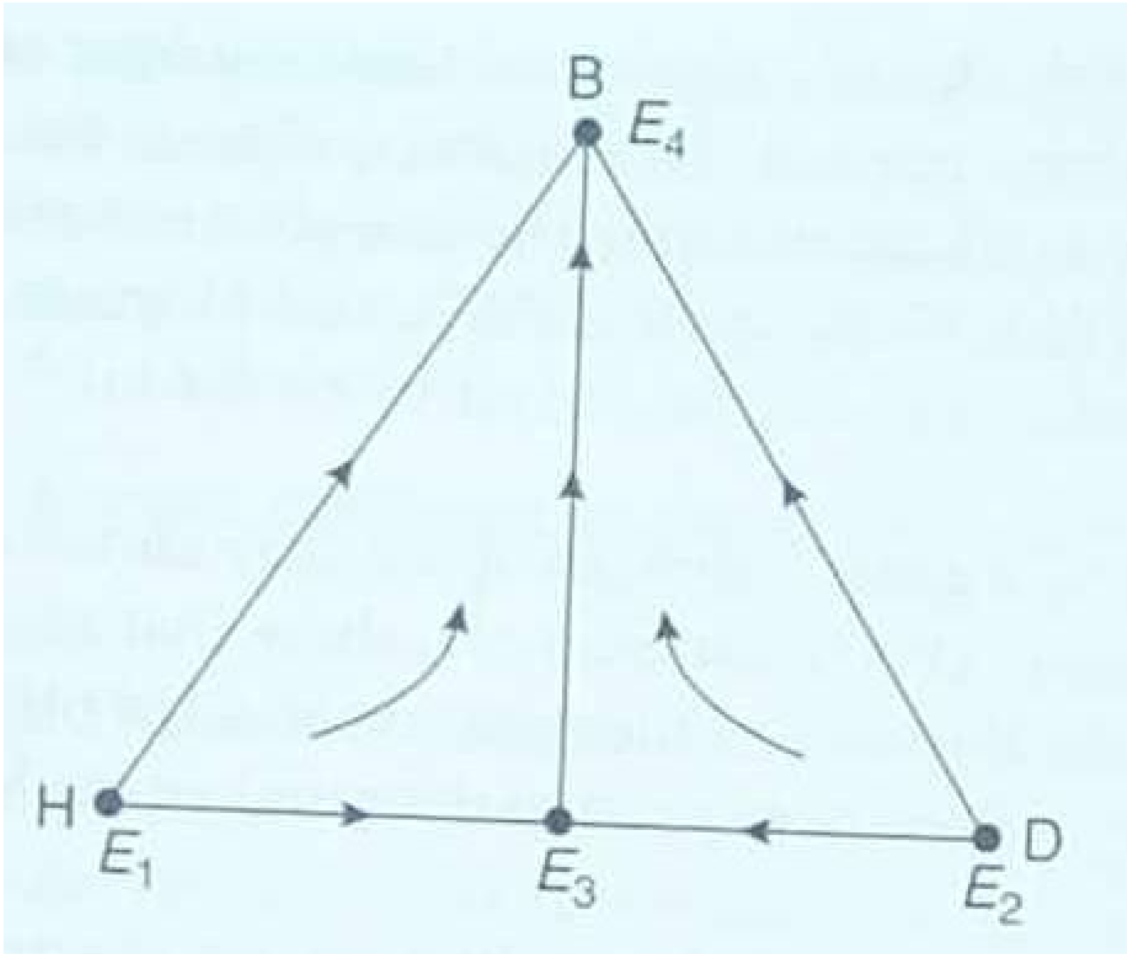

Figure 2: Evolutionary Dynamics in the Hawk-Dove- Bourgeois Game

Figure 2 shows the result of dynamics based on a function specified by Friedman that gives the rate of change of a strategy's proportion based on its payoff relative to the other two strategies. Points E_1, E_2, E_3, and E_4 are all fixed points in the sense that the proportions do not change no matter which of these points the game starts from. Only point E_4 represents an evolutionarily stable equilibrium, however, and if the game starts with any positive proportion of birds playing *Bourgeois*, the proportions tend towards E_4. The original Hawk-Dove which excluded the bourgeois strategy can be viewed as the HD line at the bottom of the triangle, and E_3 is evolutionarily stable in that restricted game.

Figure 2 also shows the importance of mutation in biological games. If the population of birds is 100 percent dove, as at E_2, it stays that way in the absence of mutation, since if there are no hawks to begin with, the fact that they would reproduce at a faster rate than doves becomes irrelevant. If, however, a bird could mutate to play *Hawk* and then pass this behavior on to his offspring, then eventually some bird would do so and the mutant strategy would be successful. The technology of mutations can be important to the ultimate equilibrium. In more complicated games than Hawk-Dove, it can matter whether mutations happen to be small, accidental shifts to strategies similar to those that are currently being played, or can be of arbitrary size, so that a superior strategy quite different from the existing strategies might be reached.

The idea of mutation is distinct from the idea of evolutionary dynamics, and it is possible to use one without the other. In economics models, a mutation would correspond to the appearance of a new action in the action set of one of the players in a game. This is one way to model innovation: not as research followed by stochastic discoveries, but as accidental learning. The modeller might specify that the discovered action becomes available to players slowly through evolutionary dynamics, or instantly, in the usual style of economics. This style of research has promise for economics, but since the technologies of dynamics and mutation are important there is a danger of simply multiplying models without reliable results unless the modeller limits himself to a narrow context and bases his technology on empirical measurements.

Notes

N5.1 Finitely repeated games and the Chainstore Paradox

- The Chainstore Paradox does not apply to all games as neatly as to Entry Deterrence and The Prisoner's Dilemma. If the one-shot game has only one Nash equilibrium, the perfect equilibrium of the finitely repeated game is unique and has that same outcome. But if the one-shot game has multiple Nash equilibria, the perfect equilibrium of the finitely repeated game can have not only the one-shot outcomes, but others besides. See Benoit & Krishna (1985), Harrington (1987), and Moreaux (1985).

- John Heywood is Bartlett's source for the term "tit-for-tat," from the French "tant pour tant."

- A realistic expansion of a game's strategy space may eliminate the Chainstore Paradox. Hirshleifer & Rasmusen (1989), for example, show that allowing the players in a multi-person finitely repeated Prisoner's Dilemma to ostracize offenders can enforce cooperation even if there are economies of scale in the number of players who cooperate and are not ostracized.

- The peculiarity of the unique Nash equilibrium for the repeated prisoner's dilemma was noticed long before Selten (1978) (see Luce & Raiffa [1957] p. 99), but the term Chainstore Paradox is now generally used for all unravelling games of this kind.

- *An* **epsilon-equilibrium** *is a strategy profile s^* such that no player has more than an ϵ incentive to deviate from his strategy given that the other players do not deviate. Formally,*

$$\forall i, \;\; \pi_i(s_i^*, s_{-i}^*) \geq \pi_i(s_i', s_{-i}^*) - \epsilon, \;\; \forall s_i' \in S_i. \tag{17}$$

  Radner (1980) has shown that cooperation can arise as an ϵ-equilibrium of the finitely repeated prisoner's dilemma. Fudenberg & Levine (1986) compare the ϵ-equilibria of finite games with the Nash equilibria of infinite games. Other concepts besides Nash can also use the ϵ- equilibrium idea.

- A general way to decide whether a mathematical result is a trick of infinity is to see if the same result is obtained as the limit of results for longer and longer finite models. Applied to games, a good criterion for picking among equilibria of an infinite game is to select one which is the limit of the equilibria for finite games as the number of periods gets longer. Fudenberg & Levine (1986) show under what conditions one can find the equilibria of infinite-horizon games by this process. For the Prisoner's Dilemma, (*Always Blame*) is the only equilibrium in all finite games, so it uniquely satisfies the criterion.

- Defining payoffs in games that last an infinite number of periods presents the problem that the total payoff is infinite for any positive payment per period. Ways to distinguish one infinite amount from another include the following.

  1 Use an **overtaking criterion**. Payoff stream π is preferred to $\tilde{\pi}$ if there is some time T^* such that for every $T \geq T^*$,

$$\sum_{t=1}^{T} \delta^t \pi_t > \sum_{t=1}^{T} \delta^t \tilde{\pi}_t.$$

  2 Specify that the discount rate is strictly positive, and use the present value. Since payments in distant periods count for less, the discounted value is finite unless the payments are growing faster than the discount rate.

3 Use the average payment per period, a tricky method since some sort of limit needs to be taken as the number of periods averaged goes to infinity.

Whatever the approach, game theorists assume that the payoff function is **additively separable** over time, which means that the total payoff is based on the sum or average, possibly discounted, of the one-shot payoffs. Macroeconomists worry about this assumption, which rules out, for example, a player whose payoff is very low if any of his one-shot payoffs dips below a certain subsistence level. The issue of separability will arise again in section 13.5 when we discuss durable monopoly.

- Ending in finite time with probability one means that the limit of the probability the game has ended by date t approaches one as t tends to infinity; the probability that the game lasts till infinity is zero. Equivalently, the expectation of the end date is finite, which it could not be were there a positive probability of an infinite length.

N5.2 Infinitely Repeated Games, Minimax Punishments, and the Folk Theorem

- Aumann (1981), Fudenberg & Maskin (1986), Fudenberg & Tirole (1991a, pp. 152-62), and Rasmusen (1992a) tell more about the Folk Theorem. The most commonly cited version of the Folk Theorem says that if conditions 1 to 3 are satisfied, then:

  Any payoff profile that strictly pareto-dominates the minimax payoff profiles in the mixed extension of an n-person one-shot game with finite action sets is the average payoff in some perfect equilibrium of the infinitely repeated game.

- The evolutionary approach can also be applied to the repeated prisoner's dilemma. Boyd & Lorberbaum (1987) show that no pure strategy, including Tit- for-Tat, is evolutionarily stable in a population-interaction version of the Prisoner's Dilemma. Hirshleifer & Martinez-Coll (1988) have found that Tit-for-Tat is no longer part of an ESS in an evolutionary prisoner's dilemma if (1) more complicated strategies have higher computation costs; or (2) sometimes a *Silence* is observed to be a *Blame* by the other player. Yet biologists have found animals playing tit-for-tat— notably the sticklebacks in Milinski (1987) who can choose whether to shirk or not in investigating predator fish.

- **Trigger strategies** of *trigger-price strategies* are an important kind of strategies for repeated games. Consider the oligopolist facing uncertain demand (as in Stigler [1964]). He cannot tell whether the low demand he observes facing him is due to Nature or to price cutting by his fellow oligopolists. Two things that could trigger him to cut his own price in retaliation are a series of periods with low demand or one period of especially low demand. Finding an optimal trigger strategy is a difficult problem (see Porter [1983a]). Trigger strategies are usually not subgame perfect unless the game is infinitely repeated, in which case they are a subset of the equilibrium strategies. Recent work has looked carefully at what trigger strategies are possible and optimal for players in infinitely repeated games; see Abreu, Pearce & Staccheti (1990). Many theorists have studied what happens when players can imperfectly observe each others' actions. For a survey, see Kandori (2002).

  Empirical work on trigger strategies includes Porter (1983b), who examines price wars between railroads in the 19th century, and Slade (1987), who concluded that price wars among gas stations in Vancouver used small punishments for small deviations rather than big punishments for big deviations.

- A macroeconomist's technical note related to the similarity of infinite games and games with a constant probability of ending is Blanchard (1979), which discusses speculative bubbles.

- In the repeated Prisoner's Dilemma, if the end date is infinite with positive probability and only one player knows it, cooperation is possible by reasoning similar to that of the Gang of Four theorem in Section 6.4.

- Any Nash equilibrium of the one-shot game is also a perfect equilibrium of the finitely or infinitely repeated game.

N5.3 Reputation: The One-Sided Prisoner's Dilemma

- *A game that is repeated an infinite number of times without discounting is called a* **supergame**. There is no connection between the terms "supergame" and "subgame."

- The terms, "one-sided" and "two-sided" prisoner's dilemma, are my inventions. Only the two-sided version is a true prisoner's dilemma according to the definition of note N1.2.

- Empirical work on reputation is scarce. One worthwhile effort is Jarrell & Peltzman (1985), which finds that product recalls inflict costs greatly in excess of the measurable direct costs of the operations. The investigations into actual business practice of Macaulay (1963) is much cited and little imitated. He notes that reputation seems to be more important than the written details of business contracts.

- **Vengeance and Gratitude.** Most models have excluded these feelings (although see Jack Hirshleifer [1987]), which can be modelled in two ways.

  1 A player's current utility from *Blame* or *Silence* depends on what the other player has played in the past; or
  2 A player's current utility depends on current actions and the other players' current utility in a way that changes with past actions of the other player.

  The two approaches are subtly different in interpretation. In (1), the joy of revenge is in the action of blaming. In (2), the joy of revenge is in the discomfiture of the other player. Especially if the players have different payoff functions, these two approaches can lead to different results.

N5.4 Product Quality in an Infinitely Repeated Game

- The Product Quality Game may also be viewed as a principal- agent model of moral hazard (see Chapter 7). The seller (an agent), takes the action of choosing quality that is unobserved by the buyer (the principal), but which affects the principal's payoff, an interpretation used in much of the Stiglitz (1987) survey of the links between quality and price.

  The intuition behind the Klein & Leffler model is similar to the explanation for high wages in the Shapiro & Stiglitz (1984) model of involuntary unemployment (section 8.1). Consumers, seeing a low price, realize that with a price that low the firm cannot resist lowering quality to make short-term profits. A large margin of profit is needed for the firm to decide on continuing to produce high quality.

- A paper related to Klein & Leffler (1981) is Shapiro (1983), which reconciles a high price with free entry by requiring that firms price under cost during the early periods to build up a reputation. If consumers believe, for example, that any firm charging a high price for any of the first five periods has produced a low quality product, but any firm charging a high price thereafter has produced high quality, then firms behave accordingly and the beliefs are confirmed. That the beliefs are self-confirming does not make them irrational; it only means that many different beliefs are rational in the many different equilibria.

- An equilibrium exists in the Product Quality model only if the entry cost F is just the right size to make n an integer in equation (8). Any of the usual assumptions to get around the integer problem could be used: allowing potential sellers to randomize between entering and staying out; assuming that for historical reasons, n firms have already entered; or assuming that firms lie on a continuum and the fixed cost is a uniform density across firms that have entered.

N5.5 Markov equilibria and overlapping generations in the game of Customer Switching Costs

- We assumed that the incumbent chooses its price first, but the alternation of incumbency remains even if we make the opposite assumption. The natural assumption is that prices are chosen simultaneously, but because of the discontinuity in the payoff function, that subgame has no equilibrium in pure strategies.

N5.6 Evolutionary Equilibrium: The Hawk-Dove Game

- Dugatkin & Reeve (1998) is an edited volume of survey articles on different applications of game theory to biology. Dawkins (1989) is a good verbal introduction to evolutionary conflict. See also Axelrod & Hamilton (1981) for a short article on biological applications of the Prisoner's Dilemma, Hines (1987) for a survey, and Maynard Smith (1982) for a book. Jack Hirshleifer (1982) compares the approaches of economists and biologists. Boyd & Richerson (1985) uses evolutionary game theory to examine cultural transmission, which has important differences from purely genetic transmission.

Problems

5.1. Overlapping Generations (see Samuelson [1958]) (medium)
There is a long sequence of players. One player is born in each period t, and he lives for periods t and $t+1$. Thus, two players are alive in any one period, a youngster and an oldster. Each player is born with one unit of chocolate, which cannot be stored. Utility is increasing in chocolate consumption, and a player is very unhappy if he consumes less than 0.3 units of chocolate in a period: the per-period utility functions are $U(C) = -1$ for $C < 0.3$ and $U(C) = C$ for $C \geq 0.3$, where C is consumption. Players can give away their chocolate, but, since chocolate is the only good, they cannot sell it. A player's action is to consume X units of chocolate as a youngster and give away $1-X$ to some oldster. Every person's actions in the previous period are common knowledge, and so can be used to condition strategies upon.

(a) If there is finite number of generations, what is the unique Nash equilibrium?

(b) If there are an infinite number of generations, what are two Pareto- ranked perfect equilibria?

(c) If there is a probability θ at the end of each period (after consumption takes place) that barbarians will invade and steal all the chocolate (leaving the civilized people with payoffs of -1 for any X), what is the highest value of θ that still allows for an equilibrium with $X = 0.5$?

5.2. Product Quality with Lawsuits (medium)
Modify the Product Quality game of section 5.4 by assuming that if the seller misrepresents his quality he must, as a result of a class-action suit, pay damages of x per unit sold, where $x \in (0, c]$ and the seller becomes liable for x at the time of sale.

(a) What is $\tilde{p}$ as a function of x, F, c, and r? Is $\tilde{p}$ greater than when $x = 0$?

(b)] What is the equilibrium output per firm? Is it greater than when $x = 0$?

(c) What is the equilibrium number of firms? Show that a rise in x has an ambiguous effect on the number of firms.

(d) If, instead of x per unit, the seller pays X to a law firm to successfully defend him, what is the incentive compatibility constraint?

5.3. Repeated Games (see Benoit & Krishna [1985]) (hard)
Players Benoit and Krishna repeat the game in Table 7 three times, with discounting:

Table 7: A Benoit-Krishna Game

		Krishna		
		Silence	*Waffle*	*Blame*
	Silence	10,10	$-1,-12$	$-1,15$
Benoit:	*Waffle*	$-12,-1$	8,8	$-1,-1$
	Blame	$15,-1$	$-1,-1$	$0,0$

Payoffs to: (Benoit, Krishna).

(a) Why is there no equilibrium in which the players play *Silence* in all three periods?

(b) Describe a perfect equilibrium in which both players pick *Silence* in the first two periods.

(c) Adapt your equilibrium to the twice-repeated game.

(d) Adapt your equilibrium to the T-repeated game.

(e) What is the greatest discount rate for which your equilibrium still works in the three-period game?

5.4. Repeated Entry Deterrence (medium)
Assume that Entry Deterrence I is repeated an infinite number of times, with a tiny discount rate and with payoffs received at the start of each period. In each period, the entrant chooses *Enter* or *Stay out*, even if he entered previously.

(a) What is a perfect equilibrium in which the entrant enters each period?

(b) Why is (*Stay out, Fight*) not a perfect equilibrium?

(c) What is a perfect equilibrium in which the entrant never enters?

(d) What is the maximum discount rate for which your strategy profile in part (c) is still an equilibrium?

5.5. The Repeated Prisoner's Dilemma (medium)
Set $P = 0$ in the general Prisoner's Dilemma in Chapter 1's Table 9, and assume that $2R > S+T$.

(a) Show that the Grim Strategy, when played by both players, is a perfect equilibrium for the infinitely repeated game. What is the maximum discount rate for which the Grim Strategy remains an equilibrium?

(b) Show that Tit-for-Tat is not a perfect equilibrium in the infinitely repeated Prisoner's Dilemma with no discounting.

5.6. Evolutionarily Stable Strategies (medium)
A population of scholars are playing the following coordination game over their two possible

conversation topics over lunch, football and economics. Let $N_t(F)$ and $N_t(E)$ be the numbers who talk football and economics in period t, and let θ be the percentage who talk football, so $\theta = \frac{N(football)}{N(football)+N(economics)}$. Government regulations requiring lunchtime attendance and stipulating the topics of conversation have maintained the values $\theta = 0.5$, $N_t(F) = 50,000$ and $N_t(E) = 50,000$ up to this year's deregulatory reform. In the future, some people may decide to go home for lunch instead, or change their conversation. Table 8 shows the payoffs.

Table 8: Evolutionarily Stable Strategies

		Scholar 2	
		Football (θ)	*Economics* $(1-\theta)$
Scholar 1	*Football* (θ)	1,1	0, 0
	Economics $(1-\theta)$	0, 0	5,5

Payoffs to: (Scholar 1, Scholar 2)

(a) There are three Nash equilibria: *(Football, Football)*, *(Economics, Economics)*, and a mixed-strategy equilibrium. What are the evolutionarily stable strategies?

(b) Let $N_t(s)$ be the number of scholars playing a particular strategy in period t and let $\pi_t(s)$ be the payoff. Devise a Markov difference equation to express the population dynamics from period to period: $N_{t+1}(s) = f(N_t(s), \pi_t(s))$. Start the system with a population of 100,000, half the scholars talking football and half talking economics. Use your dynamics to finish Table 9.

Table 9: Conversation Dynamics

t	$N_t(F)$	$N_t(E)$	θ	$\pi_t(F)$	$\pi_t(E)$
-1	50,000	50,000	0.5	0.5	2.5
0					
1					
2				.	

(c) Repeat part (b), but specifying non-Markov dynamics, in which $N_{t+1}(s) = f(N_t(s), \pi_t(s), \pi_{t-1}(s))$.

5.7. Grab the Dollar (medium)

Table 10 shows the payoffs for the simultaneous-move game of Grab the Dollar. A silver dollar is put on the table between Smith and Jones. If one grabs it, he keeps the dollar, for a payoff of 4 utils. If both grab, then neither gets the dollar, and both feel bitter. If neither grabs, each gets to keep something.

Table 10: Grab the Dollar

		Jones	
		Grab (θ)	*Wait* $(1-\theta)$
Smith:	*Grab* (θ)	$-1, -1$	4, 0
	Wait $(1-\theta)$	0, 4	1,1

Payoffs to: (Smith, Jones)

(a) What are the evolutionarily stable strategies?

(b) Suppose each player in the population is a point on a continuum, and that the initial amount of players is 1, evenly divided between *Grab* and *Wait*. Let $N_t(s)$ be the amount of players playing a particular strategy in period t and let $\pi_t(s)$ be the payoff. Let the population dynamics be $N_{t+1}(i) = (2N_t(i))\left(\frac{\pi_t(i)}{\sum_j \pi_t(j)}\right)$. Find the missing entries in Table 11.

Table 11: Grab the Dollar Dynamics

t	$N_t(G)$	$N_t(W)$	$N_t(total)$	θ	$\pi_t(G)$	$\pi_t(w)$
0	0.5	0.5	1	0.5	1.5	0.5
1						
2						

(c) Repeat part (b), but with the dynamics $N_{t+t}(s) = [1 + \frac{\pi_t(s)}{\sum_j \pi_t(j)}][2N_t(s)]$.

(d) Which three games that have appeared so far in the book resemble Grab the Dollar?

The Repeated Prisoner's Dilemma: A Classroom Game for Chapter 5

Consider the following Prisoner's Dilemma, obtained by adding 8 to each payoff in Table 2 from Chapter 1:

Table 12: The Prisoner's Dilemma

		Column	
		Silence	*Blame*
Row	*Silence*	7,7	-2, 8
	Blame	8,-2	**0,0**

Payoffs to: (Row,Column)

Students will pair up to repeat this game 10 times in the same pair. The objective is to get as high a summed, undiscounted, payoff as possible (*not* just to get a higher summed payoff than any other person in the class).