

- Crawford, Vincent and Joel Sobel (1982), "Strategic Information Transmission," *Econometrica*, 50: 1431-51.
- Dawkins, Richard (1989), *The Selfish Gene*, 2nd edition, Oxford University Press.
- Farrell, Joseph (1988), "Communication, Coordination, and Nash Equilibrium," *Economic Letters*, 27: 209-14.
- (forthcoming), "Meaning and Credibility in Cheap Talk Games," in M. Demster (ed.), *Mathematical Models in Economics*, Oxford University Press.
- Farrell, Joseph and Eric Maskin (1989), "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1(4): 327-60.
- Fershtman, Chaim and Daniel J. Seidmann (1990), "Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment," MEDS, KGSMS Discussion Paper, Northwestern University.
- Forges, Françoise (1986), "An Approach to Communication Equilibria," *Econometrica*, 54(6): 1375-86.
- Fudenberg, Drew and David M. Kreps (1988), "Learning, Experimentation and Equilibrium in Games," Working Paper, Massachusetts Institute of Technology.
- Fudenberg, Drew and David Levine (1989), "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4): 759-78.
- Fudenberg, Drew and Eric Maskin (1990), "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80(2): 274-9.
- Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989), "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1(1): 60-79.
- Gul, Faruk (1990), "Rational Strategic Behavior and the Notion of Equilibrium," Working Paper, GSB, Stanford University.
- Kahneman, Daniel and Amos Tversky (1989), "Reference Theory of Choice and Exchange," Working Paper, Department of Psychology, University of California, Berkeley.
- Kohlberg, Leon, and Jean-Francois Mertens (1986), "On the Strategic Stability of Equilibria," *Econometrica*, 54(5): 1003-38.
- Luce, Duncan R. and Howard Raiffa (1957), *Games and Decisions*, New York: John Wiley.
- Machina, Mark J. (1982), "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, 50: 277-323.
- Myerson, Roger (1986), "Multistage Games with Communication," *Econometrica*, 54(2): 323-58.
- Pearce, David (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52(4): 1029-50.
- (1987), "Renegotiation-Proof Equilibria: Collective Rationality and Interpersonal Cooperation," Cowles Foundation Discussion Paper No. 855, Yale University.
- Rabin, Matthew (1990), "A Model of Pregame Communication," mimeo, Department of Economics, University of California, Berkeley.
- Safra, Zvi, Uzi Segal, and Avia Spivak (1989), "Preference Reversal and Nonexpected Utility Behavior," *American Economic Review*, 80(4): 922-30.

CHAPTER 3

Explaining cooperation and commitment in repeated games

Drew Fudenberg¹

1 Introduction

Repeated games models have been one of the main tools for understanding the effects of long-run interactions, and in particular how long-run interactions make possible forms of trust and commitment that can be advantageous to some or all of the players. The most familiar example of this is the celebrated prisoner's dilemma, displayed in figure 3.1.

	Cooperate Defect	
Cooperate	2,2	-1,3
Defect	3,-1	0,0

Figure 3.1

When this game is played a single time, the unique equilibrium outcome is for both players to defect, but when the game is played repeatedly without a known terminal date and the players are sufficiently patient there are equilibria where both players always cooperate. This cooperative equilibrium has been used to explain observed trust and cooperation in many situations in economics and political science. Examples include oligopolists "implicitly colluding" on a monopoly price, Macaulay's (1963) observations that relations between a firm and its suppliers are often based on "reputation" and the threat of the loss of future business, and non-aggression and trade pacts between competing nation-states, as in the essays in Oye (1986).²

Closely related kinds of trust and commitment can arise in models where a single long-run player faces a sequence of short-run or myopic opponents.

Examples include Simon's (1951) explanation of non-contractual relations between a firm and its workers (recast in a game-theoretic model by Kreps (1986)), and the papers by Dybvig and Spatt (1980) and Shapiro (1982) on a firm who produces high-quality output because switching to low quality would cost it future sales. In these models, when the long-run player is sufficiently patient there is an equilibrium where it is always "trustworthy" and honors its contracts or keeps quality high.

In the applications of these models, analysts typically note that there is an equilibrium of the repeated game with the desired properties, and suppose that observed behavior will correspond to that equilibrium. In symmetric games where all players are long run, the equilibrium chosen is usually the most efficient symmetric equilibrium, while in games with a single long-run player the equilibrium chosen is the one that maximizes the various player's payoff. While this may be a fruitful way to understand the various applied situations, it raises a problem at the theoretical level, for both classes of games have many other equilibria. In particular, no matter how many times a stage game is repeated, or how patient the players are, repeated play of any of the stage game's equilibria is always an equilibrium of the repeated game. Thus, while repeated games models explain how cooperation, trust, or commitment might emerge, they do not predict that cooperation or commitment will occur.

What then is the basis for the widespread intuition that certain of the repeated-game equilibria are particularly reasonable? This chapter discusses two classes of potential explanations. Sections 2 and 3 discuss the literature on "reputation effects," which models the idea that players in a repeated game may try to develop reputations for certain kind of play. The intuition here, first explored by Kreps and Wilson (1982a), Milgrom and Roberts (1982), and Kreps, Milgrom, Roberts, and Wilson (1982), is that, if a player chooses to always play in the same way, his opponents will come to expect him to play that way in the future, and will adjust their own play accordingly. To model the possibility that players are concerned about their reputation, we suppose that there is incomplete information (Harsanyi, 1967) about each player's type, with different types expected to play in different ways. Each player's reputation is then summarized by his opponents' current beliefs about his type. For example, to model a central bank's reputation for sticking to the announced monetary policy, we assign positive prior probability to a type that will always stick to its announced policy. More generally, we suppose that there are different types associated with different kinds of play, which is equivalent to assuming that no player's "type" directly influences any other player's payoff function.³

With the reputation-effects approach, the question of why certain equilibria seem particularly plausible is then whether some or all of the

players will try to develop the reputations that are associated with particular equilibria. In general, the set of equilibrium reputations will depend on the players' prior beliefs about their opponents' types. It is comparatively easy to obtain restrictions on the set of equilibria by imposing strong restrictions on the players' prior beliefs, and to the extent that we feel comfortable imposing these restrictions, they can help us understand the mechanism supporting the "plausible" equilibria. But, if very strong restrictions on the priors are needed, the restrictions do not seem to constitute an explanation of why the plausible equilibria are plausible.

This raises the question of whether reputation effects have strong implications for a broad class of prior distributions. The case where reputation effects have the strongest general implications is where a single long-run player faces a sequence of short-run opponents, each of whom plays only once, as in the papers by Kreps and Wilson (1982a) and Milgrom and Roberts (1982) on the chain-store paradox.

In these games, there is only one player who has an incentive to maintain a reputation, so it may not be surprising that reputation effects are quite powerful: under a weak full-support distribution on the prior distribution, if the long-run player is patient, he can use reputation effects to obtain the same payoff as if he could publicly commit himself to whatever strategy he most prefers (Fudenberg and Levine, 1989, 1991). The reason is that, if the long-run player chooses to play the same action in every period, eventually his opponents will come to expect him to play that action in the future; and, since the opponents are short-run, they will then play a short-run best response to the action the long-run player has chosen. (This is imprecise – the conclusion requires either that the stage game be simultaneous-move, which the chain-store game is not, or that the extensive form have a special property explained in section 2.3.)

Another case where reputation effects might be thought to allow one player to commit himself is that of a single "large" player facing a great many long-lived but "small" opponents, since the large player's reward for a successful commitment is much greater. One reason for interest in this case of small opponents is that it may be a better description of the situation facing some government entities than the short-run player model. Whether reputation effects allow the big player to commit itself depends rather more on the fine structure of the game, as observed by Fudenberg and Kreps (1987).

When all players are long-run, as in the repeated prisoner's dilemma studied by Kreps, Milgrom, Roberts, and Wilson (1982), there is no distinguished player whose interests might be expected to dominate play, and so it would seem unlikely that reputation effects could lead to strong

general conclusions. It is true that strong results can be obtained for specific prior distributions over types. For example, in the repeated prisoner's dilemma, Kreps *et al.* found that if player 2's payoffs are known to be as in the usual complete-information case, while player 1 is either a type who always plays the strategy "tit-for-tat" or is a type with the usual payoffs, then with a sufficiently long finite horizon in every sequential equilibrium both players cooperate in almost every period. However, Fudenberg and Maskin (1986) showed that by varying the prior distribution, any feasible, individually rational payoff of the complete-information game can be obtained. This confirms the intuition that reputation effects on their own have little power when all players are long run. More recently, Aumann and Sorin (1989) have shown that reputation effects do pick out the unique Pareto optimal payoffs in games of pure coordination when the prior distributions on types is restricted in a particular way. Section 3 presents these results.

To summarize sections 2 and 3, reputation effects provide a strong foundation for the intuition that a single long-run player can obtain his commitment payoff, in the sense that this conclusion emerges for a wide range of prior distributions over types. In contrast, reputation effects on their own (i.e., without strong restrictions on the priors) do not help to explain why trust or cooperation might tend to emerge in games with several long-run players.

Several authors have tried to explain the emergence of trust and cooperation using the concept of an evolutionarily stable strategy or ESS. This work is described in section 4. Axelrod and Hamilton (1981), discussed in section 4.1, introduced this concept to the study of repeated games, and showed that evolutionary stability rules out the "always defect" in the repeated prisoner's dilemma with time-average payoffs. However, this is roughly the extent of its power: many other profiles are ESS, including some where players defect most of the time.

As we will see, the reason is that "mutant strategies" attempting to invade a population playing an inefficient strategy profile may be severely "punished" by the prevailing population for deviating from the prescribed path of play. For example, the strategy "alternate between cooperate and defect; if anyone deviates from this pattern, defect forever afterwards" is an ESS, even though it uses the non-ESS profile "always defect" to punish deviations. This suggests that, in order for ESS to restrict the set of equilibria in a repeated game, there must be some reason that the punishments for deviation will not be too severe.

Two recent papers develop this idea by introducing different forces that lead to bounds on how strong punishments can be. Fudenberg and Maskin (1990a) introduce "noise" into the model by supposing that players

sometimes make "mistakes" and play a different action than they had intended to. Since any prescribed punishment may be triggered by a mistake, certain extreme punishments are ruled out. Returning to the prisoner's dilemma, the strategy that enforced alternation between cooperate and defect with the punishment of "always defect" can be invaded by a strategy that follows the alternation until a mistake is made, but arranges to return to cooperation following a deviation from prescribed play. A more complicated argument shows that the only evolutionarily stable outcome of a pure strategy ESS is for both players to cooperate in almost every period. However, the same assumptions do not imply efficiency in other repeated games. These results are presented in section 4.2.

Section 4.3. presents the model of Binmore and Samuelson (1991), who analyze noiseless repeated games where players incur an implementation cost that depends on the number of "states" (in the sense of automata theory) required to implement their strategy. This cost implies that players will not use a strategy that has states that are not reached in the course of play, and in particular rules out strategies that punish deviations for an infinite number of periods before returning to the equilibrium path. Using this fact, Binmore and Samuelson show that ESS outcomes in any repeated game must be efficient.

2 REPUTATION EFFECTS WITH SINGLE LONG-RUN PLAYER

2.1 The chain-store game

The literature on reputation effects began with the papers by Kreps and Wilson (1982a) and Milgrom and Roberts (1982) on reputation in Selten's (1978) chain-store game. To set the stage for their work, let us first review a slight variant of Selten's original model. Each period, an entrant decides whether to enter or stay out of a particular market. If the entrant stays out, the incumbent enjoys a monopoly in that market; if the entrant enters, the incumbent must choose whether to fight or to accommodate. The incumbent's payoffs are $a > 0$ if the entrant stays out, 0 if the entrant enters and the incumbent accommodates, and -1 if the incumbent fights. The incumbent's objective is to maximize the discounted sum of its per-period payoffs; δ denotes the incumbent's discount factor. Each entrant has two possible types, tough and weak. Tough entrants always fight; a weak entrant has payoff 0 if it stays out, -1 if it enters and is fought, and $b > 0$ if it enters and the incumbent accommodates. Each entrant's type is private information, and each entrant is tough with probability q^0 independent of the others.⁴ Thus the incumbent has a short-run incentive to accommodate,

while a weak entrant will enter only if it expects the probability of fighting to be less than $b/(b+1)$. The incumbent faces a different entrant at each period t , and each entrant is informed of the actions chosen at all previous dates.

If this game has a finite horizon, there is a unique sequential equilibrium, as Selten (1978) observed: the incumbent accommodates in the last period, so the last entrant always enters, so the incumbent accommodates in the next-to-last period, and by backwards induction the incumbent always accommodates and every entrant enters. Selten called this a "paradox" because, when there are a large number of entrants and q^0 is small, the equilibrium seems counterintuitive: one suspects that the incumbent would be tempted to fight to try to deter entry, and that the "right" prediction is that the incumbent would fight and the weak entrants would stay out. This intuition is partly supported by the fact that, in the infinite-horizon version of the model, if $a(1-q^0) - q^0 > 0$, so that the incumbent's average payoff is higher when it fights than when it accommodates, and the incumbent is sufficiently patient, i.e., the discount factor δ is close to 1, there are subgame-perfect equilibria where entry is deterred.⁵ One such equilibrium is for the incumbent to fight all entrants so long as he has never accommodated in the past, and accommodate otherwise, and for the weak entrants to stay out if the incumbent has never accommodated, and enter otherwise.⁶

Since the infinite-horizon model also has an equilibrium in which every entrant enters, it does not explain why the entry deterrence equilibrium is the most plausible. To provide an explanation, Kreps and Wilson, and Milgrom and Roberts, modified the finite-horizon model to allow the incumbent to maintain a reputation for "toughness." Specifically, suppose now that the incumbent is either "tough" or "weak." If it is weak, it has the payoff function described above; if it is tough, it has "always fight" as its dominant strategy.⁷ The entrants do not know the incumbent's type; each entrant assigns the same probability p to the incumbent being tough. (Note that the incumbent's type is chosen once and for all, and influences its preferences in every market.)

If the game is played only once, the weak incumbent accommodates if entry occurs, so that a weak entrant nets $(1-p^0)b - p^0$ from entry. Thus a weak entrant enters if $p^0 < b/(b+1) = p$ and stays out if the inequality is reversed. (Here and henceforth knife-edge cases will be ignored.)

Now suppose that the incumbent will play two different entrants in succession, in two different markets. Entrant 1 is faced first, and entrant 2 observes the outcome in market 1 before making its own entry decision.

The equilibrium of this game depends on the prior probabilities and the parameters of the payoff functions; the equilibrium is unique except for parameters on the boundaries of the regions described below.

(i) If $a\delta(1-q^0) < 1$ or $q^0 > \bar{q} \equiv (a\delta - 1)/a\delta$, the maximum benefit of fighting is less than its cost. In this case, since a weak incumbent would not fight in market 1, a weak entrant 1 enters if $p^0 < \bar{p}$, and stays out if $p^0 > \bar{p}$. Entrant 2 enters if the incumbent accommodates, and stays out if the incumbent fights.

(ii) If $q^0 < \bar{q}$, the weak incumbent is willing to fight in market 1 if doing so deters entry, since accommodating reveals the incumbent is weak and causes entry to occur. The exact nature of the equilibrium again depends on the prior p that the incumbent is tough.

(ii.a) If $p^0 > \bar{p}$, then the weak incumbent fights in market 2, weak entrants stay out of both markets, and the incumbent's expected payoff is:

$$[(1-q^0)a - q^0] + \delta(1-q^0)a.$$

(ii.b) If $p^0 < \bar{p}$, then in equilibrium the weak incumbent fights in market 2 with positive probability β less than 1.⁸ Whether the weak entrant enters in market 2 depends on whether the total probability of fighting in market 2 exceeds $b/(b+1)$, which turns out to be the case if $p^0 > (b/(b+1))^2$. Note that for $p \in [(b/(b+1))^2, b/(b+1)]$ the weak incumbent's expected average payoff is positive, while its payoff was zero for the same parameters in the one-entrant game. If $p^0 < (b/(b+1))^2$, the weak entrant enters in market 2, and the weak incumbent's payoff is 0.

As the number of markets (and entrants) increases, the size of the prior p^0 required to deter entry (for $q^0 < \bar{q}$) shrinks, so that even a small amount of incomplete information can have a very large effect in long games. When $\delta = 1$, the unique equilibrium has the following form:

(a) If $q^0 > a/(a+1)$, then the weak incumbent accommodates at the first entry, which occurs (at the latest) the first time the entrant is tough. Hence, as the number of markets $N \rightarrow \infty$, the incumbent's average payoff per period goes to zero.

(b) If $q^0 < a/(a+1)$, then for every p^0 there is a number $n(p^0)$ so that, if there are more than $n(p^0)$ markets remaining, the weak incumbent's strategy is to fight with probability 1. Thus weak entrants stay out when there are more than $n(p^0)$ markets remaining, and the incumbent's average payoff per period approaches $(1-q^0)a - q^0$ as $N \rightarrow \infty$.⁹

It is easy to explain the role played by the expression $a(1-q^0) - q^0$ in the above. Imagine that the incumbent is given a choice at time zero of making an observed and enforceable commitment either to always fight or to

always acquiesce. If the incumbent always fights, its expected payoff is $a(1-q^0)-q^0$, as it must fight the tough entrants to deter the weak ones. The asymptotic nature of the equilibrium turns exactly on whether a commitment to always fight is better than a commitment to always accommodate, which yields payoff 0. Thus one interpretation of the results is that reputation effects allow the incumbent to credibly choose whichever of the two commitments it prefers.

Note though that neither of these commitments need be the one the incumbent would like most. Suppose that $a(1-q^0)-q^0 > 0$, so that a patient incumbent is willing to fight the tough entrants to deter the weak ones. Then while it receives a positive payoff from committing to always fight, it could do even better by committing to fight with probability $b/(b+1)$, which is the minimum probability of fighting that deters the weak entrants. This commitment would give it average payoff $a(1-q^0)-bq^0/(b+1)$, which exceeds the payoff of $a(1-q^0)-q^0$ from committing to always fight. Of course, if the only two types of incumbent with positive prior probability are the tough and weak types described above, then the first time the incumbent accommodates its reputation for toughness is gone and all subsequent entrants enter. The next section discusses how reputation effects may permit commitments to mixed strategies.

2.2 Reputation effects with a single long-run player

General simultaneous-move stage games

If we view reputation effects as a way of supporting the intuition that the long-run player should be able to commit himself to any strategy he desires, the chain-store example raises several questions: does the strong conclusion derived above depend on a backwards induction argument from a fixed (and thus perfectly foreseen) finite horizon, or do reputation effects have a similar impact in the infinitely repeated version of the game? Can the long-run player maintain a reputation for playing a mixed strategy when such a reputation would be desirable? How robust are the strong conclusions in the chain-store game to changes in the prior distribution to allow more possible types? How does the commitment result extend to games with different payoffs and/or different extensive forms? What if the incumbent's action is not directly observed, as in a model of moral hazard?

To answer the first question, the role of the finite horizon, consider the infinite horizon version of the chain-store game with $a(1-q^0)-q^0 > (1-\delta)/\delta$, so that, as we saw in the last section, even if $p^0 = 0$ there is an equilibrium in which the weak incumbent fights all entrants. Note now that this is not the only equilibrium of the infinite-horizon model, and that even

if $p^0 > 0$ there is an equilibrium in which the weak incumbent accommodates the first entry. Here are the equilibrium strategies: "The tough incumbent always fights. The weak incumbent accommodates to the first entry, and then fights all subsequent entry if it has not accommodated two or more times in the past. Once the incumbent acquiesces twice, it accommodates to all subsequent entry. Tough entrants always enter: weak entrants enter if there has been no previous entry or if the incumbent has already accommodated at least twice; weak entrants stay out otherwise." In this equilibrium, the weak incumbent reveals its type by accommodating in the first period; it is willing to do so because weak entrants stay out even after the incumbent's type is revealed.

These two equilibria shows that reputation effects need not determine a unique equilibrium in an infinite-horizon model. This is potentially troubling, since it raises the possibility that the power of reputation effects in the chain-store game might rely on the power of long chains of backwards induction, and several authors have argued that such chains should be viewed with suspicion.¹⁰ At the same time, note that if the incumbent is patient it does almost as well in the new equilibrium as in the equilibrium where all entry is deterred, so that the new equilibrium does not show that reputation effects have no force in infinite-horizon models. Finally, the multiplicity of equilibria suggests that it might be more convenient to try to characterize the set of equilibria without determining all of them explicitly.

This is the approach used in Fudenberg and Levine (1989, 1991). We extend the intuition developed in the chain-store example to general games where a single long-run player faces a sequence of short-run opponents. To generalize the introduction of a "tough type" in the chain-store game, we suppose that the short-run players assign positive prior probability to the long-run player being one of several different "commitment types," each of which plays a particular fixed stage-game strategy in every period. The set of commitment types thus corresponds to the set of possible "reputations" that the long-run player might maintain.

Instead of explicitly determining the set of equilibrium strategies, we obtain upper and lower bounds on the long-run player's payoff that hold in any Nash equilibrium of the game.¹¹ The (1989) paper considers reputations for pure strategies and deterministic stage games; the (1991) paper allows for reputations for playing mixed strategies, and also allows the long-run player's actions to be imperfectly observed, as in the Cukierman and Meltzer (1986) model of the reputation of a central bank when the other players observe the realized inflation rate but not the bank's action.¹²

The upper bound on the long-run player's Nash equilibrium payoff converges, as the number of periods grows and the discount factor goes to

one, to the long-run player's Stackelberg payoff, which is the most he could obtain by publicly committing himself to any of his stage-game strategies. If the short-run player's action does not influence the information that is revealed about the long-run player's choice of stage-game strategy (as in a simultaneous-move game with observed actions) the lower bound on payoffs converges to the most the long-run player can get by committing himself to any of the strategies for which the corresponding commitment type has positive prior probability. If the stage game is not simultaneous move, the lower bound must be modified, as explained in section 2.3.

Consider a single long-run player 1 facing an infinite sequence of short-run player 2s in a simultaneous-move stage game, where the long-run player's realized choice of stage-game strategy a_1 is revealed at the end of each period. The history $h^t \in H^t$ at time t then consists of past choices $(a_1, a_2)_{k=1, \dots, t-1}$. (This would not be the case in a sequential-move game, where the observed outcome need not reveal how a player would have played at some unreach information set, or in a game where actions are only imperfectly observed.) The long-run player's type $\theta \in \Theta$ is private information; θ influences player 1's payoff but has no direct influence on player 2's; θ has prior distribution p which is common knowledge. Player 1's (behavior) strategy is a sequence of maps σ_1^t from the history H^t and Θ to the space of stage-game mixed strategies $\mathcal{A}_1$: a strategy for the period- t player 2 is $\sigma_2^t: H^t \rightarrow \mathcal{A}_2$. Since the short-run players are unconcerned about future payoffs, in any equilibrium each period's choice of mixed stage-game strategy a_2 will be a best response to the anticipated marginal distribution over player 1's actions. Let $r: \mathcal{A}_1 \rightarrow \mathcal{A}_2$ be the short-run player's best response correspondence.

Two subsets of the set Θ of player 1's types are of particular interest. Types $\theta_0 \in \Theta_0$ are "sane types" whose preferences correspond to the expected discounted value of per-period payoffs $v_1(a_1, a_2, \theta_0)$. All sane types use the same discount factor δ and maximize their expected present discounted payoffs. (The chain-store papers had a single "sane type" whose probability is close to 1.) The "commitment types" are those who play the same stage-game strategy in every period; $\theta(a_1)$ is the commitment type corresponding to a_1 . The set of commitment strategies $C_1(p)$ for prior p are those for which the corresponding commitment strategies have positive prior probability. I will present the case where Θ and thus C_1 are finite; our (1989) paper considers extensions to densities over commitment types.

Define the Stackelberg payoff for $\theta_0 \in \Theta_0$ to be:

$$v_1^s(\theta_0) = \max_{a_1} \left[\max_{a_2 \in r(a_1)} v_1(a_1, a_2, \theta_0) \right],$$

and let the Stackelberg strategy for type θ_0 be one that attains this maximum.¹³ This is the highest payoff type θ_0 could obtain if (1) his type were public information and (2) he could commit himself to always play any of his stage-game actions (including mixed actions). Note that, as in the chain-store game, the Stackelberg strategy need not be pure.

Given the set of possible (static) "reputations" $C_1(p)$, we ask which reputation from this set type θ_0 would most prefer, given that the short-run players may choose the best response that the long-run player likes least. This results in payoff:

$$v_1^*(p, \theta_0) = \sup_{a_1 \in C_1(p)} \left[\min_{a_2 \in r(a_1)} v_1(a_1, a_2, \theta_0) \right],$$

which is type θ_0 's commitment payoff relative to the set of possible reputations.

The formal model allows the prior p to assign positive probability to types that play mixed strategies. Is this reasonable? Suppose that of the 100 periods to date, where entry has occurred, the incumbent has fought in fifty of them, and that various statistical tests fail to reject the hypothesis that the incumbent's play is the result of independent $1/2 - 1/2$ randomizations between fight and acquiesce. How should the entrants expect him to play? Arguably it is reasonable to suppose that they predict a $1/2$ chance of fighting, which is consistent with a prior that assigns positive probability to a type that mixed in this way.¹⁴

Let $N(\delta, p, \theta)$ and $\bar{N}(\delta, p, \theta)$, respectively, be the lowest and highest payoffs, of type θ in any Nash equilibrium of the game with discount factor δ and prior p .

THEOREM 1 (Fudenberg and Levine, 1991) Suppose that the long-run player's choice of stage game strategy a_1 is revealed at the end of each period. Then for all θ_0 with $p(\theta_0) > 0$, and all $\lambda > 0$, there exists $\bar{\delta} < 1$ such that for all $\delta \in (\bar{\delta}, 1)$:

$$(1 - \lambda)v_1^*(p, \theta_0) + \lambda \min_{a_1, a_2} v_1(a_1, a_2, \theta_0) \leq \bar{N}_0(\delta, p, \theta_0), \quad (1a)$$

and

$$\bar{N}_0(\delta, p, \theta_0) \leq (1 - \lambda)v_1^s + \lambda \max_{a_1, a_2} v_1(a_1, a_2, \theta_0). \quad (1b)$$

Remarks

- (1) The theorem says that if type θ_0 is patient he can obtain about his commitment payoff relative to the prior distribution, and that regardless of the prior a patient type cannot obtain much more than its Stackelberg

payoff. Note that the lower bound depends only on which feasible reputation type θ_0 wants to maintain and is independent of the other types that p assigns positive probability and of the relative likelihood of different types.

(2) Of course the lower bound depends on the set of possible commitment types: if no commitment types have positive prior probability, then reputation effects have no force! For a less trivial illustration, modify the chain-store game presented above by supposing that each period's entrant, in addition to being tough or weak, is one of three "sizes," large, medium, or small, and the entrant's size is public information. It is easy to specify payoffs so that the incumbent's best pure-strategy commitment is to fight the small and medium-sized entrants, and accommodate the large ones. The theorem shows that the incumbent can achieve the payoff associated with this strategy if the associated commitment type has positive prior probability. However, if as in section 2.1 the only commitment type fights all entrants regardless of size, then the incumbent cannot maintain a reputation for fighting only the small and medium entrants, for the first time it accommodates a large entrant it reveals that it is weak.

(3) For a fixed prior distribution p , the upper and lower bounds can have different limits as $\delta \rightarrow 1$. In generic¹⁵ simultaneous-move games, $v_1^*(p, \theta_0) = v_1^*(\theta_0)$ when the prior assigns a positive density to every commitment strategy.

(4) The Stackelberg payoff supposes that the short-run players correctly forecast the long-run player's stage game action. The long-run player can obtain a higher payoff if its opponents mispredict its action. For this reason, for a fixed discount factor less than 1, some types of the long-run player can have an equilibrium payoff that strictly exceeds their Stackelberg level, as the short-run players may play a best response to the equilibrium actions of other types.

For example, in the chain-store game suppose that $a(1 - q^0) < q^0 b / (b + 1)$, so that the weak incumbent's Stackelberg payoff is zero. And suppose that the prior probability of the "tough" type is greater than $b / (b + 1)$. Then one equilibrium is for the weak incumbent to always acquiesce, and the weak entrants stay out until they have seen a tough entrant enter and the incumbent acquiesce. Then the weak incumbent's equilibrium payoff is positive, since the first entrant might happen to be weak and thus stay out. However, as $\delta \rightarrow 1$ the incumbent's average discounted payoff (i.e., the discounted payoff normalized by $(1 - \delta)$) tends to 0, as the only way the incumbent can repeatedly deter entry is to fight when a tough entrant enters, and the cost of doing so is outweighed by the

benefits. A second example is in Bénabou and Laroque (1989), where an informed insider can use his information to "take advantage" of uninformed outsiders who believe that the insider might be honest. Each time the insider takes advantage, the outsiders attach a lower probability to his honesty, so that outsiders cannot be fooled by very much very often. Intuitively, stage-game payoffs above the Stackelberg level are informational rents that come from the short-run players not knowing the long-run player's type. In the long run, the short-run players cannot be repeatedly "fooled" about the long-run player's play, and the only way the long-run player can maintain a reputation for playing a particular action is to actually play that action most of the time.¹⁶ This is why a patient long-run player cannot do better than its Stackelberg payoff. Reputation effects can serve to make commitments credible, but, in the long run, this is all that they do.

(5) While the theorem is stated for the limit $\delta \rightarrow 1$ in an infinite-horizon game, the same result covers the limit as the horizon grows to infinity of finite-horizon games with time-average payoffs.

Sketch of proof I will give an overview of the general argument and a detailed sketch for the case of commitment to a pure strategy. Fix a Nash equilibrium $(\hat{\sigma}_1, \hat{\sigma}_2)$. (Recall that σ denotes a strategy profile in the repeated game, as opposed to the stage game.) This generates a probability distributed π over histories. The short-run players will use π to compute their posterior beliefs about θ at every history that π assigns positive probability. Now consider a type $\bar{\theta}$ with $p(\bar{\theta}) > 0$, and imagine that player 1 chooses to play type $\bar{\theta}$'s equilibrium strategy $\sigma_1(\cdot | \bar{\theta}, \cdot) \equiv \bar{\sigma}_1(\cdot | h')$. This generates a sequence of actions with positive probability under π .

Since the short-run players are myopic, and best response correspondences are upper hemi-continuous, Nash equilibrium requires that the short-run player's action be close to a best response to σ_1 in any period where the observed history has positive probability and the expected distribution over outcomes is close to that generated by σ . Because the short-run players have a finite number of actions in the stage game, this conclusion can be sharpened: if the expected distribution over outcomes is close to that generated by $\bar{\sigma}_1$, the short-run players must play a best response to $\bar{\sigma}_1$.

More precisely, for any h' with $\pi(h') > 0$, let $\rho(h') = \pi[\hat{\sigma}_1(\cdot | \cdot, h') = \bar{\sigma}_1(\cdot | h') | h']$.

Lemma 1: For any $\bar{\theta}$ with $p(\bar{\theta}) > 0$ there is a $\rho < 1$ such that $\hat{\sigma}_2 \in r(\hat{\sigma}_1)$ whenever $\rho(h') > \bar{\rho}$.

Conversely, in any period where the short-run players do not play a best

Note that the proof does not assert that $p(\bar{\theta}|h')$ converges to 1 when player 1 uses type $\bar{\theta}$'s strategy. This stronger assertion is not true. For example, in a pooling equilibrium where all types play the same strategy, $p(\bar{\theta}|h')$ is equal to the prior in every period. Rather the proof shows that, if player 1 always plays like type $\bar{\theta}$, eventually the short-run players become convinced that he will play like $\bar{\theta}$ in the future.

2.3 Extensive form games

Theorem 1 assumes that the long-run player's choice of stage-game strategy is revealed at the end of each period, as in a simultaneous-move game. The following example shows that the long-run player may do much less well than predicted by theorem 1 if the stage game is sequential move. This may seem surprising, because the chain-store game considered by Kreps and Wilson (1982a) and Milgrom and Roberts (1982) has sequential moves; indeed the chain-store game and the example below have the same game tree, but different payoffs.

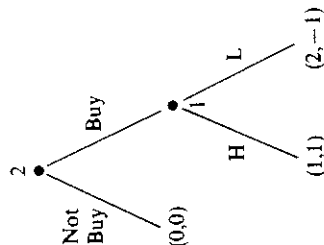


Figure 3.2

Player 2 begins by choosing whether or not to purchase a good from player 1. If he does not buy, both players receive 0. If he buys, 1 must decide whether to produce low or high quality. High quality gives each player a payoff of 1, while low quality gives player 1 a payoff of 2 and gives -1 to player 2. Note that, if player 2 does not buy, player 1's (contingent) choice of quality is not revealed. The Stackelberg outcome here is for player 1 to commit to high quality, so that all player 2s will purchase. Thus if theorem 1 extended to this game it would say that if there is positive probability p^* that player 1 is a type θ^* who always produces high quality, and if δ is close to 1, then a "sane" type θ_0 of player 1 (whose payoffs are as in the figure) receives payoff close to 1 in any Nash equilibrium.

response to $\bar{\sigma}_1$, when player 1's action is observed there is a non-negligible probability that the short-run players will be "surprised" and will increase the posterior probability that player 1 is type $\bar{\theta}$ by a non-negligible amount. After sufficiently many of these surprises, the short-run players will attach a very high probability to player 1 playing $\bar{\sigma}_1$ for the remainder of the game. In fact, one can show that for any ε there is a $K(\varepsilon)$ such that with probability $(1 - \varepsilon)$ the short-run players play best responses to $\bar{\sigma}_1$ in all but $K(\varepsilon)$ periods, and that this $K(\varepsilon)$ holds uniformly over all equilibria, all discount factors, and all priors p with the same prior probability of θ .

Given a $K(\varepsilon)$ that holds uniformly, the lower bound on payoffs is derived by considering θ to be a commitment type which has positive prior probability, and observing that type θ_0 gets at least the corresponding commitment payoff whenever the short-run players play a best response to $\bar{\sigma}_1 = \sigma(\theta)$. To obtain the upper bound, let $\bar{\theta} = \theta_0$, so that type θ_0 plays its own equilibrium strategy. Whenever the short-run players are approximately correct in their expectations about the marginal distribution over actions, type θ_0 cannot obtain much more than its Stackelberg payoff.

In general the stage-game strategies prescribed by $\bar{\sigma}_1$ may be mixed. Obtaining the bound $K(\varepsilon)$ on the number of "surprises" is particularly simple when $\bar{\sigma}_1$ prescribes a the same pure strategy $\bar{\sigma}_1$ in every period for every history. Fix an $\bar{\sigma}_1$ such that the corresponding commitment type θ has positive prior probability, and consider the strategy for play 1 of always playing $\bar{\sigma}_1$. From claim 1, there is a $\bar{p}$ such that in any period where the players 2s do not play a best response to $\bar{\sigma}_1$, $\rho(h') < \bar{p}$. Then if player 1 plays $\bar{\sigma}_1$ in every period, there can be at most $\ln(p(\theta)/\ln(\bar{p}))$ periods where this inequality obtains.

To see this, note that $\rho(h') \geq p(\bar{\theta}|h')$, because $\bar{\theta}$ always plays $\bar{\sigma}_1$. Along any history with positive probability, Bayes rule implies that:

$$p(\bar{\theta}|h^{t+1}) = p(\theta|(a^t, h^t)) = \pi(a^t|\bar{\theta}, h^t)p(\bar{\theta}|h^t)/\pi(a^t|h^t). \quad (2)$$

Then since player 2's play is independent of θ , and the choices of the two players at time t are independent conditional on h , $\pi(a^t|h^t) = \pi(a_1^t|h^t) \cdot \pi(a_2^t|h^t)$ and $\pi(a^t|\bar{\theta}, h^t) = \pi(a_1^t|\bar{\theta}, h^t)\pi(a_2^t|h^t)$. If we now consider histories where player 1 always plays $\bar{\sigma}_1$, $\pi(a_1^t|h^t) = 1$, and (2) simplifies to:

$$p(\bar{\theta}|h^{t+1}) = p(\bar{\theta}|h^t)/\pi(a_2^t|h^t). \quad (3)$$

Consequently $p(\bar{\theta}|h^{t+1})$ is non-decreasing, and increases by at least $1/\bar{p}$ whenever $\rho(h^t) \leq \bar{p}$. Thus there can be at most $\ln(p(\bar{\theta})/\ln(\bar{p}))$ periods where $\rho(h^t) \leq \bar{p}$, and the lower bound on payoffs follows. (The additional complication posed by types θ that play mixed strategies is that $p(\bar{\theta}|h^t)$ need not evolve deterministically when player 1 uses strategy σ_1 .)

This extension is false. Take $p(\theta_1) = 0.99$, $p^* = 0.01$, and consider the following strategy profile. The high-quality type always produces high quality: the "sane" type θ_0 produces low quality as long as no more than a single short-run player has ever made a purchase, and produces high quality beginning with the second time a short-run player buys. The short-run players do not buy unless a previous short-run player has already bought, in which case they buy so long as all short-run purchasers but the first have received high quality. This strategy profile is a Nash equilibrium that gives type θ_0 a payoff of 0; the profile can be combined with consistent beliefs to form a sequential equilibrium as well.¹⁷

The reason that reputation effects fail in this example is that, when the short-run players do not buy, player 1 does not have an opportunity to signal his type. This problem does not arise in the chain-store game, for there the one action the entrant can take that "hides" the incumbent's action was precisely the action the incumbent wished to be played. One response to the problem posed by the example is to assume that some consumers always purchase, so that there are no zero-probability information sets. The second response is to weaken the theorem. Let the stage-game be a finite extensive form of perfect recall without moves by Nature. As in the example, the play of the stage game need not reveal player 1's choice of normal-form strategy a_1 . However, when both players use pure strategies the information revealed about player 1's play is deterministic. Let $0(a_1, a_2)$ be the subset of A_1 corresponding to strategies a_1 , of player 1 such that (a_1, a_2) leads to the same terminal node as (a_1, a_2) . We will say that these strategies are observationally equivalent. For each a_1 let $W(a_1)$ satisfy

$$W(a_1) = \{a_2 \mid \text{for some } a_1' \text{ with support in } 0(a_1, a_2), a_2 \in r(a_1')\}. \quad (4)$$

In other words, $W(a_1)$ is the set of pure strategy best responses for player 2 to beliefs about player 1's strategy that are consistent with the information revealed when that response is played. Then, if δ is near to one, player 1's equilibrium payoff should not be much less than:

$$v_1^2(\theta_0) = \max_{a_1} \left[\min_{a_2 \in W(a_1)} v_1(a_1, a_2, \theta_0) \right]. \quad (5)$$

This is verified in Fudenberg and Levine (1989). Observe that this result, while not as strong as the assertion in theorem 1 that player 1 can pick out his preferred payoff in the graph of B , does suffice to prove that player 1 can develop a reputation for "toughness" in the sequential-move version of the chain-store game. In this game $B(\text{fight}) = \{\text{out}\}$ and $B(\text{acquiesce}) = \{\text{in}\}$. Also, $0(\text{fight}, \text{out}) = 0(\text{acquiesce}, \text{out}) = \{\text{acquiesce}, \text{fight}\}$, while $0(\text{fight}, \text{in}) = \{\text{fight}\}$ and $0(\text{acquiesce}, \text{in}) = \{\text{acquiesce}\}$. First we argue that

$W(\text{fight}) = F(\text{fight})$. To see this observe that $W(\text{fight})$ is at least as large as $B(\text{fight}) = \{\text{out}\}$. Moreover, "in" is not a best response to "fight," and "acquiesce" is not observationally equivalent to "fight" when player 2 plays "in." Consequently, no strategy placing positive weight on "in" is in $W(\text{fight})$. Finally, since player 1's Stackelberg action with observable strategies is fight, and $W(\text{fight}) = B(\text{fight})$, the generalized Stackelberg payoffs and the usual one coincide in this game.

2.4 Reputation effects with a single "big" player against many small but long-lived opponents

The previous sections showed how reputation effects allow a single long-lived player to commit itself when facing a sequence of short-run opponents. An obvious question is whether a similar result obtains for a single "big" player who faces a large number of small but long-lived opponents. For example, one might ask if a large "government" or "employer" could maintain its desired reputation against small agents whose lifetimes are of the same order as the large player's.

Suppose first that an infinite-lived "large" player faces a continuum of infinite-lived "small" players in a repeated game, that all players use the same discount factor δ , and that the small players' payoff functions are public information. The incumbent has various possible reputations, corresponding to "types" with positive probability. One might expect that the small players should behave as if their play had no influence on the play of their opponents, but this need not be the case: if the play of an individual short-run player can be observed, then there can be equilibria where its play influences the play of its opponents, even though its play has no direct effect on the payoff of any other player.¹⁸ We can however restrict attention to games where the actions of the small players are ignored by postulating that each player can only observe the play of the large player and of subsets of small players of positive measure. (Doing so provides only a starting point for the analysis, as one would like to know if there are sequences of equilibria with finitely many short-run players that converge to other limits.) Under this assumption, the small players will behave myopically. That is, each period they will play a best response to that period's expected play. Thus the situation is strategically equivalent to the case of short-run players, and theorem 1 should be expected to apply – the large player should be able to approximate the payoff of her most preferred positive-probability reputation. ("Should be expected to" because at this writing no one has worked out a careful version of the argument, attending to the technical niceties involved in continuum-of-players models.)

Next consider a large player facing a large but finite number of small long-run opponents, each of whose actions can be observed. If the payoffs of the small players are public information, one would expect that there will be equilibria which approximate the commitment result of the previous paragraph, although there could be other equilibria as well. But once the actions of the small players can be observed, they potentially have an incentive to maintain reputations of their own. Thus, instead of requiring the payoffs of the small players to be public information, it seems more reasonable to allow for the possibility that the small players will maintain reputations by supposing that there is some small prior probability that each small player is a "commitment type." The question then becomes whether the large player's concern for his reputation will outweigh the concern of the small players for theirs. Fudenberg and Kreps (1987) study this question in the context of one particular game, a multiplayer version of the two-sided concession game of Kreps and Wilson (1982a).

In the concession game, at each instant $t \in [0, 1]$, both players decide whether to "fight" or "concede." The "tough" types always fight, while the "weak" ones find fighting costly but are willing to fight to induce their opponent to concede in the future. More specifically, both weak types have a cost of 1 per unit time of fighting. If the entrant concedes first at t , the weak incumbent receives a flow of a per unit time until the end of the game, so the weak incumbent's payoff is at $-(1-t)$, and the weak entrant's payoff is $-(1-t)$. If the weak incumbent concedes first at t , the weak incumbent's payoff is $-(1-t)$ and the weak entrant's payoff is $bt - (1-t)$, where b is the entrant's flow payoff once the incumbent concedes. Thus each (weak) player would like the other player to concede, and each player will concede if he thinks his opponent is likely to fight until the end. The unique equilibrium involves the weak type of one player conceding with positive probability at date zero (so the corresponding distribution of stopping times has an "atom" at zero); if there is no concession at date zero both players concede according to smooth density functions thereafter.

In the multiplayer version, a "large" incumbent is simultaneously involved in N such concession games against N different opponents. Each entrant plays only against the incumbent, but observes play in all of the games. The incumbent is tough in all of the games with prior probability p^0 , and weak in all of them with complementary probability; each entrant is tough with prior probability q^0 independent of the others. This situation differs from that of the preceding section, in that both the big player and the small ones have the ability to maintain reputations.

The nature of the equilibrium depends on whether an entrant is allowed to resume fighting after it has already dropped out. In the "captured contests" version of the game, if an entrant has ever conceded (i.e., exited

from the market), it must concede from then on, while the "reentry" version allows the entrant to revert to fighting after it has acquiesced. Note that, when there is only one entrant, the captured contests and reentry versions have the same sequential equilibrium, as once the entrant chooses to concede it receives no subsequent information about the incumbent's type and thus will choose to acquiesce from then on.¹⁹

One might guess that if there are enough entrants the incumbent could deter entry in either version of the game. This turns out not to be the case. Specifically, under captured contests, when each entrant has the same prior probability of being tough, then no matter how many entrants the incumbent faces, equilibrium play in each market is exactly as if the incumbent played against only that entrant. To see why, suppose that there are N entrants, and that, at time t , $N-k$ of them have conceded, so that there are k entrants still fighting. Supposing that the equilibrium is symmetric (one can show that it must be) then the incumbent has the same posterior beliefs q_t about the type of each active entrant. Further supposing that the incumbent is randomizing at date t , it must be indifferent between conceding now, in which case it receives continuation payoff of zero in the remaining markets, and fighting on for a small interval dt and then conceding. The key is that whatever happens in the active markets the captured markets remained captured, so the incumbent does not consider them in making its current plans. If we denote the probability each entrant concedes between t and $t+dt$ by σ_t , we have:

$$0 = -k + k(1 - q_t)\sigma_t dt.$$

Note that the number of active entrants k factors out of this equation, so that it is the same equation as that for the one-entrant case. This is why adding more entrants has no effect on equilibrium play.

In contrast, if reentry is allowed, and there are many entrants, then reputation effects can enable the incumbent to obtain approximately its commitment payoff, provided that once the incumbent is revealed to be weak, all entrants who have previously conceded reenter.²⁰ In this case, when the incumbent has captured a number of markets he has a great deal to lose by conceding. Here the incumbent's myopic incentive is to concede to entrants who have fought a long time and thus are likely to be tough, but the incumbent lacks the flexibility to concede to these active entrants without also conceding to the entrants who have already been revealed to be weak, and this lack of flexibility enables the incumbent to commit itself to tough play. As the number of entrants grows, the equilibria converge to the profile where the incumbent never concedes and weak entrants concede immediately. At this point the remaining entrants are revealed to be tough, and the incumbent would like to concede to them if it could do so without

Thi
Wc
Bai
seri
fiel
Wh
this
mo
CO
1 A
2 A
3 E
4 R
5 h
6 h
m
d
T
M

also conceding against the weak entrants. However, since $a(1 - q^0) - q^0 > 0$, the incumbent is willing to fight the tough entrants to retain control of the other markets.

The moral of these observations is that the workings of reputation effects with one big player facing many small long-run opponents can depend on aspects of the game's structure that would be irrelevant if the small opponents were played sequentially. Thus in applications of game theory one should be wary of general assertions that reputation effects will allow a large player to make its commitments credible.

What happens when the incumbent's type need not be the same in each contest is an open question. If the types in each market are statistically independent, then the various contests can be decoupled; the interesting situation is one of imperfect correlation. One issue here is that, when types are imperfectly correlated, the incumbent's payoff aggregates outcomes in markets where it is tough and markets where it is soft, so that the exact specification of the "tough" type's payoffs and strategies becomes more important. For example, is the incumbent willing to sacrifice payoff in a market where it is weak to increase its payoff in a market where it is tough? The answer is presumably is context-specific, it might be interesting to explore some special cases.

3 REPUTATION IN GAMES WITH MANY LONG-RUN PLAYERS

3.1 General stage games and general reputations

So far we have looked at cases where reputation effects can allow a distinguished "large" or "long-run" player to commit himself to his preferred strategy. There are also incentives to maintain reputations when all players are equally large or patient, but here it is more difficult to draw general conclusions about how reputation-effects influence play.

Kreps, Milgrom, Roberts, and Wilson (1982) analyzed reputation effects in the finitely repeated prisoner's dilemma of figure 3.1. If both types are sane with probability one, then the unique Nash equilibrium of the game is for both players to defect in every period, but intuition and experimental evidence suggest that players may tend to cooperate. To explain this intuition, Kreps *et al.* introduced incomplete information about player 1's type, with player 1 either a "sane" type, or a type who plays the strategy "tit-for-tat," which is "I play today whichever action you played yesterday." They showed that, for any fixed prior probability ϵ that player 1 is tit-for-tat, there is a number K independent of the horizon length T such that, in any sequential equilibrium, both players must cooperate in almost

all periods before date K , so that, if T is sufficiently large, the equilibrium payoffs will be close to those if the players always cooperated. The point is that a sane player 1 has an incentive to maintain a reputation for being tit-for-tat, because, if player 2 were convinced that player 1 plays tit-for-tat, player 2 would cooperate until the next-to-last period of the game.

Just as in the chain-store game, adding a small amount of the right sort of incomplete information to the prisoner's dilemma yields the "intuitive" outcome as the essentially unique prediction of the model with a long finite horizon. However, unlike games with a single long-run player, the resulting equilibrium is very sensitive to the exact nature of the incomplete information specified, as was shown by Fudenberg and Maskin (1986).

Fix a two-player stage-game g with finite set of pure actions A_i for each player i and payoff functions u_1 and u_2 . Now consider repeated play of an incomplete-information game with the same action spaces as g , but where the players' payoffs need not be the same as in the repeated version of g . Call player i "sane" if his payoff is the expected value of the sum of u_i . (We can take the discount factor equal to 1 without loss of generality because we consider a large, but finite horizon.)

THEOREM 2 (Fudenberg and Maskin, 1986) For any feasible, individually rational payoff²¹ v and any $\epsilon > 0$ there exists a $\hat{T}$ such that for all $T > \hat{T}$ there exists a T -period game such that each player i has probability $(1 - \epsilon)$ of being sane, independent of the other, and such that there exists a sequential equilibrium where player i 's expected average payoff if sane is within ϵ of v_i .

Remark Note that this theorem asserts the existence of a game and of an equilibrium; it does not say that all equilibria of the game have payoffs close to v . Note also that no restrictions are placed on the form of the payoffs that players have when they are not sane, i.e., on the support of the distribution of types: no possible types are excluded, and there is no requirement that certain types have positive prior probability. However, the theorem can be strengthened to assert the existence of a game with a strict equilibrium²² where the sane types' payoffs are close to v , and a strict equilibrium of a game remains strict when additional types are added whose prior probability is sufficiently small.

Partial Proof Much of the intuition for the result can be gained from the case of payoffs v that Pareto-dominate the payoffs of a static equilibrium. Let e be a static equilibrium with payoffs $y = (y_1, y_2)$, and let v be a payoff vector that Pareto-dominates y . To avoid a discussion of public randomizations, assume that payoffs v can be attained with a pure action profile a , i.e., $g(a) = v$. Now consider a T -period game where each player i has two possible types, "sane" and "crazy," and crazy types have payoffs that make

the following strategy weakly dominant: "Play a_i as long as no deviations from a have occurred in the past, and otherwise play e_i ." Let $u_i = \max u_i$ and $\hat{u}_i = \min u_i$. Set:

$$\hat{T} > \max_i (u_i - (1 - \varepsilon)\hat{u}_i) / \varepsilon(v_i - y_i). \quad (6)$$

Consider the extensive game corresponding to $T = \hat{T}$. This game has at least one sequential equilibrium for any prior beliefs; pick one and call it the "endgame equilibrium."

Now consider $T > \hat{T}$. It will be convenient to number periods backwards, so that period T is the first one played and period 1 is the last. Consider strategies that specify that profile a is played for all $t > \hat{T}$, and that if a deviation does occur at some $t > \hat{T}$ (i.e., "before date $\hat{T}$ "), then e is played for the rest of the game, while, if a is played in every period until $\hat{T}$, play follows the endgame equilibrium corresponding to prior beliefs. Let the beliefs prescribe that, if any player deviates before $\hat{T}$, that player is believed to be sane with probability one, while if there are no such deviations before $\hat{T}$ then the beliefs are the same as the prior until $\hat{T}$ is reached, at which time strategies are given by the endgame equilibrium thereafter.

Let us check that these strategies and beliefs form a sequential equilibrium. The beliefs are clearly consistent (in the Kreps-Wilson sense). They are sequentially rational by construction in the endgame equilibrium, and are also sequentially rational in all periods following a deviation before $\hat{T}$, where both types of both players play the static equilibrium strategies.

It remains only to check that the strategies are sequentially rational along the path of play before $\hat{T}$. Pick a period $t > \hat{T}$ where there have been no deviations to date. If player i plays anything but a_i , he receives at most $\hat{u}_i$ today and at most y_i thereafter, for a continuation payoff of:

$$\hat{u}_i + (t - 1)y_i. \quad (7)$$

If instead he follows the (not necessarily optimal) strategy of playing a_i each period until his opponent deviates and playing e_i thereafter, his expected payoff will be at least:

$$\varepsilon v_i + (1 - \varepsilon) [\hat{u}_i + (t - 1)y_i]. \quad (8)$$

as this strategy yields v_i if his opponent is crazy and at least $\hat{u}_i + (t - 1)y_i$ if his opponent is sane. The definition of $\hat{T}$ has been chosen so that (8) exceeds (7) for $t > \hat{T}$, which shows that player i 's best response to player j 's strategy must involve playing a_i until $\hat{T}$. (A best response exists by standard arguments.) The key in the construction is that, when players respond to deviations as we have specified, any deviation before $\hat{T}$ gives only a one-period gain (relative to y_i). In contrast, playing a_i until $\hat{T}$ risks only a

one-period loss and gives probability ε of a gain $(v_i - y_i)$ that grows linearly in the time remaining. This is why even a very small ε makes a difference when the horizon is sufficiently long.²³

3.2 Common interest games and bounded-recall reputations

Aumann and Sorin (1989) consider reputation effects in the repeated play of two-player stage games of "common interests," which they define as stage games where there is a payoff vector that strongly Pareto dominates all other feasible payoffs. In these games, the Pareto-dominant payoff vector corresponds to a static Nash equilibrium, but there can be others, as in the game of figure 3.3.

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

Figure 3.3

This is the game used by Aumann (1990) to argue that even a unique Pareto-optimal payoff need not be the inevitable result of preplay negotiation: Player 1 should play D if he assesses probability greater than $\frac{1}{8}$ that player 2 will play R. Also, player 1 would like player 2 to play L regardless of how player 1 intends to play. Thus, when the players meet each will try to convince the other that they will play their first strategy, but these statements need not be compelling.

Aumann and Sorin show that when the possible reputations (i.e., crazy types) are all "pure strategies with bounded recall" (to be defined shortly) then reputation effects pick out the Pareto-dominant outcome so long as only pure-strategy equilibria are considered. A pure strategy for player i has recall k if it depends only on the last k choices of his opponents, that is if all histories where i 's opponent has played the same actions in the last k periods induce the same action by player i . (Note that, when player i plays a pure strategy and does not contemplate deviations, conditioning on his own past moves is redundant.) When k is large the bounded recall condition may seem innocuous, but it does rule out some simple but "unforgiving" strategies, such as those that prescribe permanent reversion to a static Nash equilibrium if any player ever deviates.

Aumann and Sorin consider perturbed games with independent types, where each player's type is private information, each player's payoff

assumption with the remark that "in a culture in which irrational people have long memories, rational people are less likely to cooperate." Note that the theorem concerns the case where δ is large compared to the recall length μ , while one might expect that a more patient player would tend to have a longer memory. This is important for the proof: if μ grew with δ , it is not clear that player 2 would try to learn player 1's strategy.

3.3 Reputation effects in repeated bargaining models with two long-run players

Schmidt (1990) extends the logic of the proof of Fudenberg and Levine (1989) to a repeated bargaining model where both players are long-lived but only one of the players has the opportunity to build a reputation. In this model, in each period t a seller whose cost of production is known (and equal to 0 without loss of generality) makes an offer p_t to a buyer whose value v is private information; v takes on finitely many values between $\bar{v}$ and $\underline{v}$. If the buyer accepts offer p_t , the seller's payoff that period is p_t and the buyer's payoff is $v - p_t$; if the buyer rejects each player's payoff that period is zero. The good is not storable, and each period the seller has a new good to offer for sale. Each player's objective is to maximize the expected discounted sum of his per-period payoffs, with discount factor δ_b for the buyer and δ_s for the seller. (Hart and Tirole (1988) solved this game for the case where the buyer and seller have the same discount factor and the buyer has only two possible types.)

Since the seller's payoff function is assumed to be public information, only the buyer potentially has the ability to develop a reputation. If we can show that the buyer always rejects prices above his valuation, then each valuation is a "commitment type" in the sense of section 2.2, and the buyer's most preferred reputation is for $v = \bar{v}$. (This reputation only yields the Stackelberg payoff if $\bar{v}$ equals the seller's cost of zero.) While it seems intuitive that the buyer should behave in this way, the infinite-horizon model has equilibria where this is not the case. However, Schmidt shows that the buyer does reject all prices above his valuation in a Markov-perfect equilibrium of the finite-horizon model, so that $\bar{v}$ does serve as a commitment type when this equilibrium concept is used.²⁵ The strategy "reject all prices above $\bar{v}$ " is the corresponding commitment strategy, and the seller's best response to this strategy is to always charge price $\bar{v}$.

The proof of theorem 1 relies on the fact that short-run players always play a short-run best response to the anticipated play of their opponents, so the short-run players would necessarily be "surprised" in any period where their play was not a best response to the long-run player's play. In the context of the bargaining model, this says that if the seller were a short-run

function depends only on his own type, and types are independently distributed. The prior p_i about player i 's type is that player i is either the "sane" type θ_0 , with the same payoffs as in the original game, or a type that plays a pure strategy with recall less than some bounded μ . Moreover, p_i is required to assign positive probability to the types corresponding to each pure strategy of recall 0. These types play the same action in every period regardless of the history, just like the commitment types considered in section 2.²⁴ Such priors correspond to "admissible perturbations of recall μ ," or " μ -perturbations" for short. Say that a sequence p of μ -perturbations supports a game G if $p^m(\theta_0) \rightarrow 1$ for all i where $m \rightarrow \infty$, and if the conditional distribution $p^m(\theta^i | \theta^i \neq \theta_0)$ is constant.

THEOREM 3 (Aumann-Sorin, 1989) Let the stage-game g be a game of common interests, and let z be its unique Pareto-optimal outcome. Fix a recall length μ , and let p^m be a sequence of μ -perturbations that support the associated discounted repeated game $G(\delta)$. Then the set of pure-strategy Nash equilibria of the games $G(\delta, p^m)$ is not empty, and the pure-strategy equilibrium payoff converge to z for any sequence (δ, m) converging to $(1, \infty)$.

Idea of Proof The intuition is clearest in the case where δ goes to 1 much faster than m goes to infinity (the theorem holds uniformly over sequences (δ, m)). Suppose a pure-strategy equilibrium exists, and suppose its payoff is less than z . Consider the strategy for player 1 of always playing the action $a_1(z)$ corresponding to z . Since the equilibrium is pure this strategy is certain to eventually reveal that player 1 is not type θ_0 . The commitment type $\theta(z)$ corresponding to $a_1(z)$ has positive probability by assumption, so if $\mu = 0$ player 2 will infer that player 1 is $\theta(z)$ and will play $a_1(z)$ from then on (because crazy types play constant strategies when $\mu = 0$). However, player 1 could be some other type with memory longer than 0, and to learn player 1's type will require player 2 to "experiment" to see how player 1 responds to different actions. Such experiments could be very costly if they provoked an unrelenting punishment by player 1, but since player 1's crazy types all have recall at most μ , player 2's potential loss (in normalized payoff) from experimentation goes to zero as δ goes to one. Thus if δ is sufficiently large we expect player 2 to eventually learn that player 1 has adopted the strategy "always play $a_1(z)$," and so when δ is close to one player 1 can obtain approximately z by always playing $a_1(z)$.

Remarks Aumann-Sorin give counterexamples to show that the assumptions of bounded recall and full support on recall 0 are necessary, and also show that there can be mixed-strategy equilibria whose payoffs are bounded away from z . They interpret the necessity of the bounded recall

player he would be "surprised" whenever his offer was rejected. Because the seller is a long-run player, it is not necessarily true that he will never make an offer that is certain to be refused, and so there can be periods in which the seller does not play a best response to the buyer's commitment strategy and yet is not surprised when the commitment strategy is played. Nevertheless, as Schmidt shows, even a long-run seller cannot continually make offers that are likely to be refused, as the seller could do better by charging price $\bar{v}$, which is certain to be accepted in any Markov perfect equilibrium. More precisely, for any discount factor $\delta_s < 1$, and any $\varepsilon > 0$, there is an $M(\varepsilon, \delta_s)$ such that among the first M offers above $\bar{v}$, at least one of them has a probability of acceptance of at least ε . With this result in hand, the proof of theorem 1 can be used to conclude that, in any Markov perfect equilibrium, if the buyer adopts the strategy of always rejecting prices above $\bar{v}$, then eventually the seller must charge $\bar{v}$. This implies that as $\delta_b \rightarrow 1$ and the horizon $T \rightarrow \infty$ for a fixed δ_s , the buyer's equilibrium payoff converges to its commitment value of $\bar{v} - \bar{v}$. (The seller's discount factor must be held fixed, as $M(\varepsilon, \delta_s)$ goes to infinity as δ_s goes to 1.)

Moreover, because of the assumption that the game has a finite horizon, this conclusion can be strengthened to obtain for all $\delta_b > 1/2$. Schmidt shows that $\delta_b > 1/2$ implies that the few "bad" periods where the seller's price exceeds $\bar{v}$ occur towards the end of the game. (More precisely, there is a K independent of the length of the game T such that the seller offers price $\bar{v}$ whenever there are at least $T - K$ periods remaining.) Thus with a sufficiently long horizon even an impatient buyer obtains approximately his commitment payoff.²⁶

4 EVOLUTIONARY STABILITY IN REPEATED GAMES

While the idea of applying evolutionary stability to repeated games is roughly as old as the literature on reputation effects, so far it has not been as extensively developed, and it will receive correspondingly less attention in this chapter.

4.1 An introduction to evolutionary stability in repeated games

Consider a symmetric two-player game, meaning that both players have the same sets S and Σ of feasible pure and mixed strategies, respectively, and the same utility function $u(\cdot, \cdot)$, where the first argument is the strategy chosen by the player and the second argument is the strategy of the player's opponent.

A strategy profile σ in a symmetric two-player game is a "strictly evolutionarily stable strategy" or "strict ESS" (Maynard Smith and Price,

1973; Maynard Smith, 1974) if no other strategy profile σ' has as high a payoff as σ against the strategy $(1 - q)\sigma + q\sigma'$ for all sufficiently small positive q . When the space of pure strategies is finite, this condition is equivalent to the condition that for all $\sigma' \neq \sigma$, either:

- $u(\sigma', \sigma) < u(\sigma, \sigma)$, or (i)
- $u(\sigma', \sigma) = u(\sigma, \sigma)$ and $u(\sigma, \sigma') > u(\sigma', \sigma')$.²⁷ (ii)

A weak ESS is a profile σ such that for every $\sigma' \neq \sigma$ satisfies either (i) or the weaker condition (ii').

- $u(\sigma', \sigma) = u(\sigma, \sigma)$ and $u(\sigma, \sigma') \geq u(\sigma', \sigma')$. (ii')

This definition allows σ to repel invasion by σ' by doing as well as σ' against the mixtures of σ and σ' .²⁸ Inspection of (i) and (ii') makes it clear that an evolutionary stable profile is a symmetric Nash equilibrium; the second clause in (ii) gives evolutionary stability additional bite. The intuition for the concept is that, if s is not evolutionarily stable, it can be invaded by a "mutant" strategy s' : if a small percentage of a large group of players begins to play s' , and players are randomly matched with a different opponent each period, then the expected payoff of s' exceeds that of s , and this may mean that the percentage of players using s' will increase.

In the biological justification of the concept, it is supposed that the strategy each individual plays is determined by its genes, and that individuals reproduce copies of themselves at a rate proportional to their payoff in the game. Moreover, it is supposed that all of the animals belong to the same population, as opposed to there being distinct populations of "player 1s" and "player 2s." (Actually the usual biological model leads not to ESS but rather to something called the "replicator dynamics": the fraction of the population playing strategy s grows at a rate proportional to the difference between the payoff to using s and the average payoff obtained in the whole population. Note that this dynamics is deterministic, and does not allow for "mutations." An evolutionary stable profile is a stable fixed point of the replicator dynamics, but other profiles can be stable as well.²⁹)

Even in animal populations, it is not clear that the "hard-wired" interpretation of the determinants of behavior should be taken literally, as behavior may be thought to be coded for by several genes that co-evolve in a complex way. Nevertheless, theoretical biologists have found evolutionary stability a useful concept for explaining animal behavior.

When applied to human agents, the hard-wired interpretation is even more controversial. Instead, the assumption that the growth rates of the population fractions using each strategy are proportional to the strategies' payoffs has been defended as the result of various kinds of boundedly rational learning (see, e.g., Sugden, 1986; Crawford, 1989). For example,

each period a small proportion of the population might somehow learn the current payoff of each strategy, and choose the strategy with the highest current payoff. A more appealing story might be that players learn the strategies and current payoffs of a few other individuals ("neighbors?") and again myopically choose the strategy with the highest current payoff. However, some other learning processes do not lead to concepts like evolutionary stability, and there is not yet much of a consensus on which economic contexts evolutionary stability is appropriate for. The interest of the results reported below relies on the hope that either a good foundation will be found for the application of evolutionary stability to economics, or that the results will turn out to extend to related equilibrium concepts for which economic foundations can be provided.

The first application of evolutionary stability to repeated games was by Axelrod and Hamilton (1981). They showed that the strategy "always defect" is not evolutionarily stable in the repeated prisoner's dilemma with time average payoffs.³⁰ In particular, they noted that a population using "always defect" can be invaded by the strategy "tit-for-tat," which cooperates in the first period and then always plays the strategy its opponent played the period before. Tit-for-tat can invade because it does as well against always defect as always defect does against itself (both give a time-average payoff of 0) and tit-for-tat obtains payoff 2 when paired against itself, so that tit-for-tat does strictly better than always defect for any proportion q of mutants playing tit-for-tat.

If players discount their repeated-game payoffs with discount factor δ , this conclusion needs to be modified, as the payoff of tit-for-tat against always defect is then not 0 but $-(1-\delta)$. In this case, tit-for-tat cannot invade if its proportion q is arbitrarily small and strategies are randomly matched with each other, as the probability $(1-q)$ of losing $(1-\delta)$ outweighs the potential gain of $2q$. However, tit-for-tat can still invade if there is a sufficient amount of "clustering," meaning that mutants are matched with each other more often than random matching would predict. (Clustering is discussed by Hamilton (1964); with the payoffs of figure 3.1 it suffices that the probability that tit-for-tat is paired with itself be greater than $(1-\delta)/(3-\delta)$.)

Thus evolutionary stability can be used to rule out the strategy "always defect." Axelrod and Hamilton argued further that evolutionary stability supports the prediction that players will use the strategy tit-for-tat. They noted that tit-for-tat is evolutionarily stable with time-average payoffs, and indeed is evolutionarily stable for discount factors sufficiently close to 1, even if clustering is allowed.³¹ They also noted that tit-for-tat was the winning strategy in two computer tournaments organized by Axelrod (1980a, 1980b) (entrants in the second tournament were informed of the

results in the first one) and that tit-for-tat eventually dominated play when the strategies submitted in Axelrod (1980b) were allowed to evolve according to the replicator dynamics.

Although the experimental results are interesting, the theoretical argument is weak. The problem is that many strategies besides tit-for-tat are evolutionarily stable with time-average payoffs. In particular, the outcome where players always defect can be approximated arbitrarily closely by an ESS. Consider the strategy "cooperate in period 0, k , $2k$, etc., and defect in all other periods, so long as the past play of both players has conformed to this pattern. If in some past period play did not conform, then defect in all subsequent periods." Call this strategy "1C, k D." This strategy yields payoff $2/k$ when matched against itself, which is close to the payoff of always defect if k is large. Yet the strategy is evolutionarily stable, because if an invader deviates from the pattern, it is punished forever afterwards and so obtains a time-average payoff of at most 0.

Note that the ESS strategy 1C, k D uses "always D " as a punishment for deviations, even though always D is not itself an ESS. A mutant strategy cannot invade by conforming to 1C, k D on the equilibrium path and improving on "always D " in the punishment states following deviations, since so long as both players conform to the equilibrium path the punishment states have probability zero. However, the fact that always D is not an ESS does suggest that 1C, k D can be invaded if players sometimes make "mistakes," so that the punishment of always D is triggered with positive probability. This observation is the starting point of the work described in the next subsection.

4.2 Evolutionary stability in noisy repeated games

Fudenberg and Maskin (1990a) use the assumption that players make mistakes (and other assumptions detailed below) to show that players always cooperate in any symmetric ESS of the repeated prisoner's dilemma with time-average payoffs. More generally, we obtained lower bounds on the payoffs in ESS of symmetric two-player stage games. Whether or not these bounds imply efficiency depends on whether there is a unique feasible payoff in the stage game where the sum of the two player's payoffs is maximized.

To begin, it would be helpful to give a precise definition of what is meant by symmetry in this context. Suppose that the stage-game g is symmetric, so that in the stage game both players have the same set A of feasible actions.³² Then the time- t history h of a player is the sequences of past actions chosen by himself and by his opponent, and a pure strategy s is a sequence of maps from histories to actions. (Note that, with this definition of the history, a

given sequence of actions generates two distinct histories, corresponding to the viewpoints of the two players.) A symmetric profile is a profile in which the two players use the same strategy. For example, the profile where both play tit-for-tat is symmetric. Symmetry does not require that the two players choose identical actions in every subgame: if both play tit-for-tat, then in the subgame following the first-period actions (C, D) , the second-period actions will be (D, C) .

Next, assume that players use strategies of only finite complexity in the sense of Kalai and Stanford (1983).³³ This means the following: say that histories h^T and h^T are *equivalent under s* if for any T and any sequence of action profiles a^T of length T , strategy s prescribes the same action following (h^T, a^T) and following $(\tilde{h}^T, a^T)$. The complexity of s is the number of distinct equivalence classes it induces. For example, the strategy tit-for-tat has two equivalence classes, one consisting of all histories where the opponent played D last period, and the other the union of the initial history and any history where the opponent played C last period. For any initial history h^T , the play of a profile of finitely complex strategies will eventually follow a repetitive cycle, so time-average payoffs are well-defined.

Finally, suppose that the game has a very small amount of "noise." "Noise" means that the realized actions are sometimes not the ones that the players intended to choose, and that each player observes only the actions his opponent actually played, and not the intended ones. The noise is small in the sense that the most likely event, with probability almost 1, is that each player never makes a "mistake." The next most likely event, with probability $\varepsilon \approx 0$, is for exactly one mistake somewhere along the infinite course of play. Each player is equally likely to make this mistake, and it can occur in any period. Two mistakes has probability about ε^2 , and so on. By taking the limit $\varepsilon \rightarrow 0$ we have a situation where the preferences of the players are lexicographic: payoffs conditional on no mistakes are infinitely more important than payoffs conditional on one mistake, which are infinitely more important than payoffs conditional on two mistakes, and so on.³⁴ If we step back from the various limits to consider the case of discount factors $\delta < 1$ and error probabilities $\varepsilon > 0$, the lexicographic preferences describe a situation where the next mistake is unlikely to happen for such a long time that its effect on payoff is negligible, so that the error probability must be tending to 0 "faster" than the discount factor tends to 1.

Say that a payoff vector $(v, v') \in V$ is efficient if it maximizes the sum of the players' payoffs, i.e., $(v, v') \in \arg \max (u + u' | (u, u') \in V)$. This definition gives equal weights to both players, which is natural given we have supposed that the "player 1s" and the "player 2s" are drawn from a common population. Let $\hat{u} = \min\{u | \text{there exists a } u' \text{ such that } (u, u') \in V \text{ is efficient}\}$. If in some subgame a player's payoff is below $\hat{u}$, not only is there an alternative

outcome of that subgame where both players are better off, but any efficient outcome *must* be better for both players. In contrast, when a player's payoff exceeds $\hat{u}$, there are efficient outcomes that he likes less. In the prisoner's dilemma this point is moot, because there is a unique efficient payoff, namely (2,2). However, in the game in figure 3.4, any feasible payoff vector that sums to 5 is efficient, and $\hat{u} = 1$. For this reason the following theorem has very different implications in these two games.

THEOREM 4 (Fudenberg and Maskin, 1990a) If (v, v) is feasible and individually rational,³⁵ and $v > \hat{u}$, and there is *some* finitely complex symmetric profile with payoffs (v, v) , then there is a finitely complex pure-strategy ESS with these payoffs. Conversely, each player's payoff in a finitely complex pure strategy ESS is at least $\hat{u}$.³⁶

Here is a partial intuition for this result. Consider first how and why some payoff vectors can be supported by ESS. In the prisoner's dilemma of figure 3.1, the unique efficient payoff is (2,2). To see that (2,2) is an ESS of the noisy repeated game, consider the profile in which both players use the strategy "perfect tit-for-tat," which is "cooperate in the first period; in subsequent periods cooperate if and only if in the previous period either both players cooperated or both players defected." Denote this strategy by σ^* . If both players use σ^* , the continuation payoffs in every subgame are (2,2): each mistake triggers one period of mutual punishment, and then cooperation resumes. At the same time, the one period of mutual punishment is enough to deter deviations, so the profile is subgame-perfect.

If perfect tit-for-tat were not an ESS, then for all q there would need to be a σ' , such that $u(\sigma', \sigma') - u(\sigma^*, \sigma') > [(1-q)/q][u(\sigma^*, \sigma^*) - u(\sigma', \sigma^*)]$. But for $u(\sigma', \sigma')$ to be close to $u(\sigma^*, \sigma^*)$, σ' must induce σ^* to cooperate in every period, and so σ' must also cooperate in almost every period. Thus $u(\sigma^*, \sigma')$ will be close to 2, and σ' cannot achieve a higher payoff when matched with itself, so perfect tit-for-tat is an ESS.³⁷

It is interesting to note that the "usual" tit-for-tat of the evolutionary biology literature—"cooperate at the start and thereafter play the action the opponent played last period"—is not a Nash equilibrium (and *a fortiori* is not an ESS) in the noisy prisoner's dilemma: the first mistake triggers an inefficient cycle with payoffs $(3, -1)$, $(-1, 3)$, etc. For the same reason tit-for-tat is not subgame-perfect in the model without noise.

The next example shows how an ESS can have inefficient payoffs.

In the game of figure 3.4, the payoff (2,2) is inefficient, but gives each player more than $\hat{u} = 1$. This payoff vector is the outcome of the ESS profile where players use the strategy: "play d in the first period, and continue to play d so long as last period either both players played d or neither did. If one player unilaterally deviates from d , henceforth he plays b and his

	a	b	c	d
a	0,0	4,1	0,0	0,0
b	1,4	0,0	0,0	0,0
c	0,0	0,0	0,0	0,0
d	0,0	0,0	0,0	2,2

Figure 3.4

opponent plays a." Even though this profile is inefficient, any strategy that tries to promote greater efficiency will be punished with payoff 1 forever; this punishment is consistent with ESS because it is an efficient static equilibrium.

To see the idea behind the converse direction of the theorem, that payoffs below $\hat{u}$ cannot be supported by an ESS, consider a finitely complex pure-strategy profile (s,s) that is "supersymmetric," meaning that both players choose the same actions in every subgame, even those where the past play of the players has been different. And suppose there is an action a^* such that (a^*,a^*) is efficient. From the assumption of finite complexity, there is some history (or histories) h' , where the players' continuation payoffs $v(h)$ are minimized. Now consider a "mutant" strategy s' that plays like s except following history h' . Given history h' , s' plays some action $\hat{a} \neq s(h')$ at period t . If its opponent plays a as well, it is revealed to be a "fellow mutant," and henceforth s' plays a^* . If its opponent does not play a at h' , at all subsequent dates s' plays just like s . Since h' is a history where continuation payoffs are minimized, when s' is paired with s it receives the same payoff as s does when paired with itself. And when s' is paired with itself, it does strictly better, since every history has positive probability of being reached.

This argument shows that a "supersymmetric" ESS not only has payoffs above $\hat{u}$ but must be efficient! When general strategies (such as tit-for-tat) are allowed, the continuation payoffs of the players can be different in histories where they have played differently in the past. This is why equilibrium payoffs need only exceed $\hat{u}$.

The restriction to pure strategies makes it easy for a "mutant" to signal its identity; we believe that this restriction is not essential. The restriction to finite complexity is essential. Consider the following strategy for the prisoner's dilemma: "Alternate between C and D so long as past play conforms to this pattern. If there is a deviation, switch to playing C every

third period, after a subsequent deviation switch to C every fourth period, and so on." Both players using this strategy is an ESS because, regardless of the history, any deviation is punished by a positive amount, but the strategy is not efficient.

Maskin and I believe we will be able to extend theorem 4 to the case of discount factors close to 1 and a 'small but non-infinitesimal' probability of mistake. Since the lexicographic model describes a situation where payoffs conditional on even one mistake are not very important, the model corresponds to a limit where the probability of mistake per period goes to zero much faster than the discount factor goes to 1. And from the discussion of the discounting case in the previous subsection, we know that evolutionary stability will only have bite if "clustering" is allowed. Our hoped-for extension will assert that, for any positive period amount q of clustering for discount factors δ close enough to 1, and error probabilities that are sufficiently small compared to $1 - \delta$, ESS exist, and ESS must have payoffs at least $\hat{u} - \epsilon$, with ϵ going to zero as q tends to 1.

We are fairly confident that this extension is true, and that it holds without the restrictions to pure strategies and finite complexity. A potentially more difficult extension is to the case of games played by two separate populations, as opposed to the single population assumed above. After all, it does not seem reasonable to assume that players are assigned each period to be either a "consumer" or a "firm." Allowing for distinct populations would also permit the analysis of ESS in repeated games where some of the players are long-lived and the others play only once, as in Fudenberg, Kreps, and Maskin (1990).

4.3 Evolutionary stability in games played by finite automata

Binmore and Samuelson (1991) consider evolutionary stability in repeated games in which less complex strategies are less costly, as in Abreu and Rubinstein (1989). In this model, players choose finite automata (idealized computer programs) to play the repeated game for them, and the cost of the automata is increasing in their complexity, which is the number of internal states the automata requires.³⁸ Players prefer strategies which maximize their "direct" payoff in the repeated game, but between two strategies with the same direct payoff, players prefer the one whose complexity is lower. Thus complexity here represents a cost of implementing the strategies, as opposed to a cost of computing payoffs or of finding the best response to a strategy of the opponent.

In a Nash equilibrium of the automata game, neither player's machine can have any states that are not used on the path of play, as such unused states could be dropped without reducing the player's direct payoff. For

example, both players playing tit-for-tat is not a Nash equilibrium, as the unique best response to tit-for-tat in the presence of implementation costs is the strategy "always cooperate." However, there are Nash equilibria that do result in players cooperating in every period but the first one. It is also a Nash equilibrium for both players to always defect.

The fact that in equilibrium every state of the automata must be reached rules out infinite punishments. Binmore and Samuelson exploit this restriction to show that all pure-strategy ESS profiles must be efficient.

THEOREM 5 (Binmore and Samuelson, 1991)³⁹ In a symmetric two-player automata game with complexity costs, every pure-strategy ESS has efficient payoffs, and ESS exist.

Sketch of Proof While Binmore and Samuelson discuss symmetrized versions of underlying asymmetric games, in which players are randomly assigned to the roles "player 1" and "player 2" each time they are matched, the proof is easier in the case where the underlying game is symmetric and the players cannot use their labels to correlate their play. Here it is clear that efficient Nash equilibria are evolutionarily stable.

To see why other Nash equilibria are not ESS, note first that in any equilibrium both players will choose the same actions in every period both on and off of the equilibrium path. That is, all pure-strategy Nash equilibria must be "supersymmetric" in the sense discussed in the proof of theorem 4. Suppose also that there is a finite automaton s^* such that (s^*, s^*) has efficient payoffs.⁴⁰

Now consider an ESS s that is not efficient, and consider a mutant strategy $\hat{s}$ that plays as follows. In the initial period, $\hat{s}$ plays an action $\hat{a}$ that differs from the initial action played by s . If the opponent's initial action is $\hat{a}$, $\hat{s}$ plays like the efficient automaton s^* from the second period on. If the opponent's initial action is not $\hat{a}$, $\hat{s}$ computes the state q that automaton s will be in following the initial actions, and plays like s in state q from the next period on. (Because strategy s must be supersymmetric, it specifies the same play for both players even after a unilateral deviation by one of them.) Since state q is certain to be reached when s plays against itself, the strategy $\hat{s}$ obtains the same average payoff when matched with strategy s as s does when matched with itself. Hence $\hat{s}$ can invade s , so s is not an ESS after all.

Note that theorem 5 asserts that every ESS of the game in figure 3.4 must be efficient, and so "always d " is not the outcome of an ESS of the automata game, although it is the outcome of an ESS of the game with noise. The reason is that, without noise, a mutant strategy that attains the efficient average payoff of $2\frac{1}{2}$ when matched against itself can only be repelled by a strategy with an infinite number of periods of punishment, and such infinite punishments are ruled out by implementation costs.

The differing conclusions of the implementation-cost model and the model with mistakes leads to the question of their relative merits. The former model describes a world in which the cost of additional states is high compared to the probability of mistakes, while the latter describes a world in which the probability of mistakes is high compared to the cost of additional states. My own view is that the latter world is a more realistic reduced form, because I think that in most populations of players there is enough randomness in the play of the next opponent that players will not be tempted to switch from tit-for-tat to always cooperate in order to save on internal states.⁴¹ However, the main source of the variation need not be mistakes. Binmore and Samuelson consider "polymorphic" populations in which different types of machines coexist. Such populations, which are analogous to mixed-strategy equilibria, permit more states to be retained, and show that mistakes are not necessary to obtain more intuitive conclusions.

5 CONCLUSIONS

This chapter began by asking how to explain the widespread intuition that certain equilibria of repeated games are particularly likely. In games with a single long-run player, the idea of reputation effects provides a strong foundation for the intuition that the long-run player should be able to obtain his commitment payoff. In games with several equally patient players, reputation effects have strong predictions only if strong restrictions are imposed on the players' prior beliefs. Evolutionary stability may provide an explanation of why long-run players tend to cooperate, but the results require assumptions about the relative magnitude of mutation probabilities and the patience of the players, and the validity of the ESS concept in economic applications is not yet resolved.

There are several other potentially interesting approaches to repeated games, but so far none of them have been able to explain either cooperation when all players are long-run, or commitment by a single long-run player. One approach is to apply equilibrium refinements based on forward induction, such as the "strategic stability" of Kohlberg and Mertens (1986). So far, this concept has only been applied to finitely repeated games, where its results can conflict with efficiency, even where efficiency is a perfect-equilibrium outcome. This makes it seem unlikely that strategic stability would explain cooperation in the infinitely repeated prisoner's dilemma, but verifying this requires an extension of the stability concept to infinite games. However, given such an extension, it seems likely that stability does predict that a single patient player can achieve his commitment payoff against a series of short-run opponents.

models, and poses the following question: in the prisoner's dilemma, why should player 1 be satisfied with teaching his opponent that he plays tit-for-tat, when a higher payoff can be obtained by teaching him that he must allow player 1 to defect occasionally without being punished? A possible answer is that it is harder to teach this "greedy" strategy, but this seems hard to formalize.

Notes

- 1 It will be clear that David Kreps, David Levine, and Eric Maskin played a major role in the development of the results reported here, but this does not fully reflect their contribution to my own understanding of the field. I would like to thank all three of them, and also Jean Tirole, for many helpful conversations. The discussion of reputation effects here draws heavily on chapter 9 of Fudenberg and Tirole (1991).
- 2 Recent economic applications of repeated games to explain trust and cooperation include Greif (1989), Milgrom, North, and Weingast (1989), Porter (1983), and Rotemberg and Saloner (1986).
- 3 This is a narrower meaning of reputation than that suggested by common usage. For example, one might speak of a worker having a "reputation" for high productivity in Spence's (1974) signaling model, and of the high-productivity workers investing in this reputation by choosing high levels of education.
- 4 This presentation of the chain-store game is based on the summary by Fudenberg and Kreps (1987). Kreps and Wilson consider only the case $q^0 = 0$, while Milgrom and Roberts consider a richer specification of payoffs.
- 5 This was observed by Milgrom and Roberts (1982).
- 6 This is an equilibrium if the incumbent's discount factor δ satisfies $a(1 - q^0) - q^0 > \frac{(1 - \delta)}{\delta}$.

- 7 By this we mean either that the tough incumbent is unable to accommodate, as in Milgrom and Roberts, or that the incumbent's payoff in the repeated game is such that all strategies but "always fight" are weakly dominated. Fudenberg, Kreps, and Levine (1988) give an algorithm for determining payoff functions with this property.
- 8 This probability is determined by the requirement that, if the incumbent fights in market 2, the posterior probability that it is tough makes the next entrant indifferent between fighting and staying out. (Recall that the weak incumbent will accommodate in market 1.)
- 9 Note that we fix p^0 and take the limit as $N \rightarrow \infty$. For fixed N and sufficiently small p^0 , the weak incumbent must accommodate in each market in any sequential equilibrium. I believe that the characterization extends to any δ by replacing the term $a/(a+1)$ with $[a - (1 - \delta)/\delta]/(a+1)$, but I have not checked the details.
- 10 Following Rosenthal (1981), this point has been made in various ways by Reny (1985), Basu (1985), and Fudenberg, Kreps, and Levine (1988).

A second approach is taken in the literature on "renegotiation" in repeated games surveyed by Pearce in this volume. This literature supposes that equilibrium is the result of negotiation between the players, and assumes that in a static game these negotiations will lead to an equilibrium that is Pareto-efficient in the set of all equilibrium payoffs. If these negotiations can only occur before play has begun, then for sufficiently large discount factors negotiations would lead to efficient outcomes. However, the literature supposes that players can meet and "renegotiate" their original (non-binding) agreement at the beginning of each period, so that equilibria with low continuation payoffs might be overturned. The question is then whether efficiency can be attained in the presence of the renegotiation constraints.

The idea of modeling players as automata is another way to try to obtain sharper predictions in repeated games. The Abreu and Rubinstein (1989) model does yield some predictions, but does not imply that outcomes must be efficient without the addition of the ESS concept. This implementation-cost literature is still in its early stages, and a correspondingly large amount of work remains before its usefulness is established. In particular, while strategies that require an infinite number of states seem implausible, it is less clear that an n -state machine is more costly than a machine with $(n-1)$ states, and that this cost difference is larger than the small probability of mistakes or other random factors. For example, Banks and Sundaram (1989) have shown that, under a measure of complexity that sums the number of states and its number of transition paths, the only Nash equilibrium of the prisoner's dilemma is to always defect. This measure of the cost of a strategy, like counting states, seems to be based on a "hardware" interpretation of complexity costs. To keep with the computer science analogy, it would be interesting to explore complexity measures based on the cost of writing the software to implement the strategy.

Finally, all of the papers I have discussed have used some equilibrium concept as a starting point. An alternative is to explicitly model the process by which equilibrium might be reached. For example, one could consider boundedly rational automata trying to "learn" their opponents' strategies, which would place the complexity cost at the level of the players' calculations instead of at the level of implementation. This approach brings the "automata" literature much closer to that literature on when equilibrium in games can be explained as the result of players having learned one another's strategies, as in Fudenberg and Kreps (1988).

From the viewpoint of learning models, one conjecture is that if the players realize that each is trying to learn the other's strategy, then each player will try to "teach" its opponent in a way that leads the opponent to play "nicely." This is reminiscent of ideas from the reputation-effects

assumption is innocuous in infinitely repeated games with little discounting (Fudenberg and Maskin, 1990b) it may not be innocuous here. In the absence of public randomizations, and if (as assumed throughout this chapter) only the realized actions in the stage game are observed, and not the players' intended randomizations, theorem 2 has only been proved for payoffs such that $v_i > \min_{a_{-i}} u_i(a_i, a_{-i})$, i.e., the minimization must be restricted to pure strategies for player i 's opponents.

22 An equilibrium is strict if each player's strategy is a strict best response to the strategies of his opponents, i.e., no other strategy does as well.

23 Note once again that as $\varepsilon \rightarrow 0$ the required $\hat{T} \rightarrow \infty$, or conversely that for a fixed T a sufficiently small ε has no effect.

24 In games between long-run players, it can be advantageous to commit to a history-dependent strategy, such as tit-for-tat in the prisoner's dilemma. In contrast, a single long-run player facing a sequence of short-run opponents can obtain the commitment payoff using a strategy of recall 0.

25 See Maskin and Tirole (1989) for a discussion of this equilibrium concept. The Markov-perfect assumption is not needed if the buyer has only two possible types; it is not known whether it is needed with three types or more.

26 The reason that theorem 1 only yields the commitment payoff in the limit as the long-run player's discount factor tends to 1 is that it covers both finite and infinite horizon games, and with an infinite horizon the "bad" periods can occur at the start of play, as in the equilibrium of the chain-store game where the first entrant is not deterred.

27 With an infinite strategy space conditions (i) or (ii) are necessary but not sufficient for the desired inequality to hold over *all* σ' for a given q .

28 Maynard Smith calls this "neutral evolutionary stability."

29 See van Damme (1987) for an excellent survey of the relationship between evolutionary stability and stability in the replicator dynamics. Boylan (1990) proposes a more general dynamics that allows for mutations.

30 In applying evolutionary stability to infinitely repeated games, one supposes that in each "round" players are paired to play the entire repeated game; after the round is over the population fractions of each strategy are updated according to its relative payoff. To allow each round to end in a finite amount of time, we can suppose that the periods of round 1 take place on the interval $[0, 1]$, round 2 takes place from $t = 1$ until $t = 2$, and so on.

An interesting alternative would be for players to reproduce each period, with a stationary probability per period that the current match is broken off and the players are rematched with others in the population.

31 Any strategy that is never the first to defect will always cooperate against tit-for-tat, so that tit-for-tat is weakly stable. It is not strictly stable, and indeed no strictly stable strategy profile exists. (Boyd and Lorberbaum (1987) prove this for pure strategies, Farrell and Ware (1989) for mixed strategies with finite support, and Y. G. Kim (1989) for general mixed strategies.) Sugden (1986) and Boyd (1989) show that strict ESS exist in the discounted repeated prisoner's

11 Recall that the set of Nash equilibria is robust to the introduction of additional types whose prior probability is small, while the set of sequential equilibria is not (Fudenberg, Kreps, and Levine, 1988).

12 Other models of reputation with imperfectly observed actions include Bénabou and Laroque (1989) and Diamond (1989).

13 Because r has a closed graph, the maxima in this definition are attained.

14 For those who are uncomfortable with the idea of types who "like" to play mixed strategies, an equivalent model identifies a countable set of types with each mixed strategy of the incumbent. Thus, one type always plays fight, the next acquiesces the first period and fights in all others, another fights every other opportunity, and so on — one type for every sequence of fight and acquiesce. Thus every type plays a deterministic strategy, and by suitable choosing the relative probabilities of the types the aggregate distribution induced by all of the types will be the same as that of the given mixed strategy.

15 Genericity is needed to ensure that, by a small change in α_i , player 1 can always "break ties" in the right direction in the definition of $v_i^*(p, \beta_0)$.

16 Hal Varian has suggested that this be called the "Abe Lincoln theorem," because it shows that the long-run player can't fool all of its opponents all of the time.

17 These strategies are not a sequential equilibrium if the horizon is finite. They thus do not form a counterexample to the sequential equilibrium version of theorem 1 for finite horizon games. Indeed, Y. S. Kim (1990) has shown that, when this game is played with a long but finite horizon, there is a unique sequential equilibrium, and in it the firm does maintain a reputation for high quality. Kim is working on the question of the best lower bound for sequential equilibria in finite repetitions of general stage games with reputation effects.

18 This point is made in Fudenberg and Levine (1988), who show that such equilibria are not artifacts of the continuum-of-players model, but rather can arise as limits of equilibria of games with a finite number of players. It may be that the common intuition that the play of a small player should be ignored corresponds to a continuum-of-players models with a noise term that masks the actions of individual players yet vanishes in the continuum-of-players limit.

19 If there are several entrants and the incumbent plays them in succession, so that $t \in [0, 1]$ is against the first entrant, $t \in [1, 2]$ against the second, and so on, then the first entrant might regret having acquiesced if it sees the incumbent acquiesce to a subsequent entrant, but at that point the first entrant's contest is over, and once again the captured contests and reentry versions have the same equilibrium.

20 Backwards induction implies that this is the unique sequential equilibrium in the discrete-time, sequential-move, finite-horizon version of the game. However, the continuous-time formulation has another equilibrium in which the entrants do not reenter.

21 A payoff vector is strictly individually rational if $v_i > \min_{a_{-i}} \max_{a_i} u_i(a_i, a_{-i})$ for all players i . Fudenberg and Maskin assumed that each period players jointly observe the outcome of a "public randomization," e.g., a "sunspot." While that

dilemma if players make "mistakes." They consider a discounted formulation with "clustering" (defined below), and so their model has a large set of ESS.

The non-existence of strict ESS is a general property of games with non-trivial extensive forms, as it cannot be satisfied by a profile that leaves some information sets unreached. This led Selten (1983, 1988) to define a "limit ESS" as the limit of a sequence of strictly evolutionarily stable strategies in "perturbed games" where players tremble and play all actions with positive probability. Selten's purpose of introducing these "mistakes," like that of Boyd and Sugden, is to *enlarge* the set of evolutionarily stable strategies to avoid non-existence problems, so he defines the limit ESS to include all strict ESS of the unperturbed game. In the work discussed in the next subsection, mistakes are used to restrict the set of (weak) ESS.

32 If the stage game is asymmetric, it can be made symmetric by assuming that at the start of each period nature randomly assigns players to one of the two roles. Then a stage-game action is a contingent map, specifying how to play in each role. The key assumption is that all players are equally likely to play each role, which is not a good description of many economic situations. Also, with this symmetrization process a number of mixed strategies yield the same behavior strategy and are thus equivalent, as noted by Selten (1983), who proposed the notion of a "direct ESS" to get around the resulting non-existence of strictly stable mixed profiles.

33 We believe that the assumption of finite complexity is unnecessary when considering discounting instead of time-averaging.

34 This is a special case of the lexicographic preferences in Blume, Brandenburger and Dekel (1990). Note that if there is an i.i.d. probability ε of mistake in each period, then for any $\varepsilon > 0$ there is probability 1 of an infinite number of mistakes, while in our model in infinite number of mistakes has probability 0.

35 As defined in note 20.

36 The published proof is incomplete: it establishes the existence of a history h that minimizes the player's continuation payoff, but implicitly assumes that this h^* is asymmetric. However, the case of asymmetric h^* is easily handled.

37 So is "perfect n -tit-for-tat," where each mistake or deviation triggers n periods of "both defect." Because we use time average payoffs, no strict ESS exists even though the model has noise. We believe that when we extend our analysis to the limit of discount factors tending to 1 we will be able to construct strict ESS.

38 This complexity measure is similar but not identical to the measure introduced by Kalai and Stanford that was discussed in the last subsection.

39 Like Fudenberg and Maskin, Binmore and Samuelson sidestep the non-existence of strict ESS by using the weak version of the concept. They also extend their result to the limit of discount factors tending to 1 as clustering probabilities tend to 0.

40 This is not the case in the game of figure 3.4, where efficiency requires asymmetric play. To implement this asymmetry requires some way of distinguishing between the players. One way to do this is to assign the players labels. Another is to introduce a probability of mistakes, or to consider mixed strategies, so that symmetric profiles can generate asymmetric histories. Under

any of these alternatives, the equilibria need no longer be supersymmetric.

41 Kalai and Neme (1989) have shown that any individually rational payoffs can be Nash equilibria if there is positive probability of even a single mistake.

References

- Abreu, D. and A. Rubinstein (1989), "The Structure of Nash Equilibrium in Repeated Games with Finite Automata," *Econometrica*, 56: 1259-81.
- Aumann, R. (1990), "Preplay Communication Does Not Lead to Nash Equilibrium," mimeo, Hebrew University.
- Aumann, R. and L. Shapley (1976), "Long-term Competition - A Game Theoretic Analysis," mimeo.
- Aumann, R. and S. Sorin (1989), "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1: 5-39.
- Axelrod, R. (1980a), "Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24: 3-25.
- (1980b), "More Effective Choice in the Prisoner's Dilemma," *Journal of Conflict Resolution*, 24: 379-403.
- (1981), "The Emergence of Cooperation Among Egoists," *American Political Science Review*, 75: 306-18.
- Axelrod, R. and W. D. Hamilton (1981), "The Evolution of Cooperation," *Science*, 211: 1390-6.
- Banks, J. and R. Sundaram (1989), "Repeated Games, Finite Automata, and Complexity," Working Paper 183, University of Rochester.
- Benabou, R. and G. Laroque (1988), "Using Privileged Information to Manipulate Markets," INSEE Working Paper 137/930.
- Blume, L., A. Brandenburger, and E. Dekel (1990), "Lexicographic Beliefs and Equilibrium Refinements," *Econometrica*, 59: 81-98.
- Boylan, R. (1990), "Equilibria Resistant to Mutation," mimeo, Caltech.
- Chamberlin, E. (1929), "Duopoly: Value Where Sellers are Few," *Quarterly Journal of Economics*, 43: 63-100.
- Crawford, V. (1989), "Learning and mixed-strategy Equilibria in Evolutionary Games," *Journal of Theoretical Biology* 140: 537-50.
- Cukierman, A. and A. Meltzer (1986), "A Theory of Ambiguity, Credibility and Inflation Under Discretion and Asymmetric Information," *Econometrica*, 54: 1099-121.
- Diamond, D. (1989), "Reputation in Acquisition and Debt Markets," *Journal of Political Economy*, August, 828-62.
- Dybvig, P. and C. Spatt (1980), "Does it Pay to Maintain a Reputation?" mimeo, Carnegie-Mellon University.
- Fudenberg, D. and D. M. Kreps (1987), "Reputation and Simultaneous Opponents," *Review of Economic Studies*, 54: 541-68.
- (1988), "A Theory of Learning, Experimentation and Equilibrium in Games," mimeo, Stanford Graduate School of Business.
- Fudenberg, D., D. M. Kreps, and D. K. Levine (1988), "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44: 354-80.

- Fudenberg, D., D. M. Kreps, and E. Maskin (1990), "Repeated Games with Long-Run and Short-Run Players," *Review of Economic Studies*, 57: 555-73.
- Fudenberg, D. and D. K. Levine (1988), "Open-Loop and Closed-Loop Equilibria in Dynamic Games with Many Players," *Journal of Economic Theory*, 44: 1-18.
- (1989), "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57: 759-78.
- (forthcoming), "Maintaining a Reputation When Strategies are Imperfectly Observed," *Review of Economic Studies*.
- Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54: 533-54.
- (1990a), "Evolution and Cooperation in Noisy Repeated Games," *American Economic Review*, 80: 274-79.
- (forthcoming), "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*.
- Fudenberg, D. and J. Tirole (1991) *Game Theory*, Cambridge, MA: MIT Press.
- Greif, A. (1989), "Reputation and Coalitions in Medieval Trade: Evidence from the Geniza Documents," mimeo, Stanford University.
- Hamilton, W. D. (1963), "The Evolution of Altruistic Behavior," *American Naturalist*, 97: 354-6.
- Harsanyi, J. (1967), "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 4: 159-82, 320-34.
- Hart, O. and J. Tirole (1988), "Contract Renegotiation and Coasian Dynamics," *Review of Economic Studies*, 55: 509-40.
- Kalai, E. and A. Neme (1989), "The Strength of a Little Perfection," mimeo, Northwestern University.
- Kalai, E. and W. Stanford (1988), "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56: 397-410.
- Kim, Y. G. (1989), "Evolutionary Stable Strategies in the Repeated Prisoner's Dilemma Game," mimeo, UCSD.
- Kim, Y. S. (1990), "Characterization and Properties of Reputation Effects in Finitely-Repeated Extensive Form Games," mimeo, UCLA.
- Kohlberg, E. and J.-F. Mertens (1986), "On the Strategic Stability of Equilibrium," *Econometrica*, 54: 1003-37.
- Kreps, D. M. (1986), "Corporate Culture and Economic Theory," in M. Tsuchiya (ed.), *Technological Innovation and Business Strategy* (Tokyo: Nippon Keizai Shimbunsha Press, in Japanese). Forthcoming in English in Alt and Shepsle (eds.), *Rational Perspectives on Political Science*, Cambridge, MA: Harvard University Press.
- Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27: 245-52, 486-502.
- Kreps, D. and R. Wilson (1982a), "Reputation and Imperfect Information," *Journal of Economic Theory*, 27: 253-79.
- (1982b), "Sequential Equilibria," *Econometrica*, 50: 863-94.

- Macaulay, S. (1963), "Non-Contractual Relations in Business: A Preliminary Study," *American Sociological Review*, 28: 55-67.
- Maskin, E. and J. Tirole (1989), "Markov Equilibrium," mimeo, Harvard and MIT.
- Maynard Smith, J. (1974), "The Theory of Games and the Evolution of Animal Conflict," *Journal of Theoretical Biology*, 47: 209-21.
- (1982), *Evolution and the Theory of Games*, Cambridge University Press.
- Maynard Smith, J. and G. R. Price (1973), "The Logic of Animal Conflicts," *Nature*, 246: 15-18.
- Milgrom, P., D. North, and B. Weingast (1989), "The Role of Law Merchants in the Revival of Trade: A Theoretical Analysis," mimeo, Stanford University.
- Milgrom, P. and J. Roberts (1982), "Predation, Reputation and Entry Deterrence," *Econometrica*, 50: 443-60.
- Neyman, A. (1986), "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma," *Economics Letters*, 19: 227-9.
- Oye, K. (1986), *Competition Under Anarchy*, Princeton, NJ: Princeton University Press.
- Pearce, D. (1991), "Repeated Games: Cooperation and Rationality," this volume.
- Rosenthal, R. (1981), "Games of Perfect Information, Predatory Pricing, and the Chain-Store Paradox," *Journal of Economic Theory*, 25: 92-100.
- Rotemberg, J. and G. Saloner (1986), "A Supergame-Theoretic Model of Price Wars During Booms," *American Economic Review*, 76: 390-407.
- Schelling, T. C. (1960), *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.
- Scherer, F. M. (1980), *Industrial Market Structure and Economic Performance*, (2nd edition), Boston: Houghton-Miffling.
- Schmidt, K. (1990), "Commitment through Incomplete Information in a Simple Repeated Bargaining Model," mimeo, University of Bonn.
- Selten, R. (1978), "The Chain-Store Paradox," *Theory and Decision*, 9: 127-59.
- (1983), "Evolutionary Stability in Extensive Two-Person Games," *Mathematical Social Sciences*, 5: 269-363.
- Shapiro, C. (1982), "Consumer Information, Product Quality, and Seller Reputation," *Bell Journal of Economics*, 13: 20-35.
- Shubik, M. (1970), "Game Theory, Behavior, and the Paradox of Prisoner's Dilemma: Three Solutions," *Journal of Conflict Resolution*, 14: 181-94.
- Simon, H. (1951), "A Formal Theory of the Employment Relationship," *Econometrica*, 19: 293-305.
- Spence, A. M. (1974), *Market Signaling*, Cambridge, MA: Harvard University Press.
- Sugden, R. (1986), *The Economics of Rights, Co-operation, and Welfare*, New York, NY: Basil Blackwell.