Reputation and Equilibrium Selection in Games with a Patient Player

Drew Fudenberg; David K. Levine

Econometrica, Vol. 57, No. 4 (Jul., 1989), 759-778.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28198907%2957%3A4%3C759%3ARAESIG%3E2.0.CO%3B2-8

Econometrica, Vol. 57, No. 4 (July, 1989), 759–778

REPUTATION AND EQUILIBRIUM SELECTION IN GAMES WITH A PATIENT PLAYER

BY DREW FUDENBERG AND DAVID K. LEVINE[1]

A single long-run player plays a simultaneous-move stage game against a sequence of opponents who play only once, but observe all previous play. Let the "Stackelberg strategy" be the pure strategy to which the long-run player would most like to commit himself. If there is positive prior probability that the long-run player will always play the Stackelberg strategy, then his payoff in any Nash equilibrium exceeds a bound that converges to the Stackelberg payoff as his discount factor approaches one. When the stage game is not simultaneous move, this result must be modified to account for the possibility that distinct strategies of the long-run player are observationally equivalent.

KEYWORDS: Reputation, inference, Stackelberg equilibrium.

1. INTRODUCTION

CONSIDER A GAME in which a single long-run player faces an infinite sequence of opponents, each of whom plays only once, but who observes all previous play. While such a game will often have multiple equilibria, a common intuition is that the "most reasonable" equilibrium is the one which the long-run player most prefers. This paper shows that "reputation effects" provide a foundation for that intuition, and it also identifies an important way in which the intuition must be qualified.

More specifically, imagine that players move simultaneously in each period. Let the "Stackelberg outcome" be the long-run player's most preferred pure strategy profile of the stage game under the constraint that the short-run player chooses a best response to the long-run player's strategy. The "Stackelberg strategy" is the corresponding strategy of the long-run player. Now suppose that with nonzero prior probability the long-run player is a "type" who always plays his Stackelberg strategy. When the discount factor is sufficiently near to one, any Nash equilibrium must give the long-run player almost his Stackelberg payoff. This result depends on the assumption that the short-run players observe the long-run player's strategy in each stage game. As we will see, the result must be modified in sequential-move games where, for example, some actions by the short-run player may preclude the long-run player from acting at all.

Our work builds on that of several previous authors, most directly that of Kreps-Wilson (1982), Milgrom-Roberts (1982), and Fudenberg-Kreps (1987) on reputation effects in the chain-store paradox. These papers considered a long-lived incumbent facing a succession of short-lived entrants, and showed that if there was a small chance that the incumbent was "tough," it could deter entry by

[1] This paper benefited from conversations with Matt McCarter, Paul Milgrom, John Riley, Jean Tirole, an editor, and two referees. National Science Foundation Grants SES 85-09697, SES 85-09484, SES 87-08616, the Sloan Foundation and UCLA Academic Senate provided financial support. Part of this research was conducted while the second author was at the University of Minnesota.

maintaining a reputation for toughness. Our result improves on theirs in several ways, all of which stem from the fact that our results apply to all of the Nash equilibria of the repeated game. First, our results are robust to further small changes in the information structure of the game. The earlier arguments depend on the restriction to sequential equilibria which, as Fudenberg-Kreps-Levine (1988) have argued, is not robust to such changes.

Second, our proof is much simpler, and provides a clearer understanding of the reputation-effects phenomenon. The key observation is that a feasible strategy for the long-run player is to always play the Stackelberg strategy. This need not be the long-run player's optimal strategy, but the optimal strategy must give him at least as high a payoff. Consequently, our technique is to find a lower bound on the payoff to always playing Stackelberg that is uniform across all of the Nash equilibria of the game. The bound results from the fact that since the short-run players are myopic they play as Stackelberg followers in any period in which they attach a high probability to the long-run player playing his Stackelberg strategy. This fact implies that if the long-run player chooses his Stackelberg strategy in every period, there is an upper bound on the number of times the short-run players can fail to play as Stackelberg followers. This argument is much simpler than the earlier ones, which obtained bounds on the long-run player's payoffs by characterizing the set of sequential equilibrium strategies.

Third, because our proof is simpler, we are able to study a broader class of games. We consider arbitrary specifications of the stage game, as opposed to the special case of the chain store, and we consider a more general form of incomplete information: Where the earlier papers specified that the long-run player had two or three types, our result covers all games in which the "Stackelberg type" has positive probability. In addition, our results extend to non-stationary games and to games where the long-run player has private information about his payoffs in addition to knowing whether or not he is a "Stackelberg type."

Our work on games with a single patient player is also related to that of Kreps-Milgrom-Roberts-Wilson (1982) and Fudenberg-Maskin (1986) on reputation effects in games where all of the players are long-lived, and, more closely, to that of Aumann-Sorin (1987). Kreps-Milgrom-Roberts-Wilson considered the effects of a specific sort of incomplete information in the finitely-repeated prisoner's dilemma, and showed that in all sequential equilibria the players cooperated in all but a few of the periods. Fudenberg-Maskin showed that for any given individually rational payoffs of a repeated game there is a family of "nearby" games of incomplete information each of which has a sequential equilibrium which approximates the given payoffs. Our result does not apply to these games because long-run players balance short-run losses against long-run gains. Consequently, unlike the short-run players we consider, long-run players need not play a short-run best response to the anticipated play of their opponents. The work of Aumann-Sorin is closer to ours in developing bounds on payoffs that hold uniformly whenever the incomplete information puts probability on a sufficiently broad class of preferences. To obtain this sort of result with several long-run players, Aumann-Sorin require very strong assumptions: The only stage games considered are those of "common interest," and the preferences of the "crazy" types must be represented by "finite automata with finite memory."

2. THE SIMPLE MODEL

We begin with the simplest model of a long-run player facing a sequence of opponents. The long-run player, player one, faces an infinite sequence of different short-lived player two's. Each period, player one selects a strategy from his strategy set S_1, while that period's player two selects a strategy from S_2. In this section we assume that players one and two move simultaneously in each period, and that at the end of the period each player learns what strategy his opponent used during that period. Section 5 considers the complications that arise when the stage game has a nontrivial extensive form. For the time being we will also assume that the S_i are finite sets; Section 7 considers the technical issues involved when the S_i are allowed to be any compact metric spaces. Corresponding to the strategy spaces S_i are the space Σ_i of mixed strategies.

The *unperturbed stage game* is a map

$$g: S_1 \times S_2 \rightarrow R^2,$$

which gives each player i's payoff g_i as a function of the realized actions. In an abuse of notation, we let $g(\sigma) = g(\sigma_1, \sigma_2)$ denote the expected payoff corresponding to the mixed strategy σ.

In the *unperturbed repeated game* $G(\delta)$, the long-run player maximizes the normalized discounted value of expected payoffs using the discount factor δ, $0 \leq \delta < 1$. Specifically, the sequence of payoffs $g_1^1, \ldots, g_1^t \ldots$ has normalized present value

$$(1-\delta) \sum_{t=0}^{\infty} \delta^t g_1^t.$$

The normalizing factor $1-\delta$ is chosen so that payoffs in the repeated game are measured in the same units as those of the stage game. Each period's short-run player acts to maximize that period's payoff.

Both long-run and short-run players can condition their play at time t on the entire past history of the game.[2] Let H_t denote the set of possible histories of the game through time t, $H_t = (S_1 \times S_2)^t$. A pure strategy for player one is a sequence of maps s_1^t: $H_{t-1} \rightarrow S_1$, and a pure strategy for a period t player two is a function s_2^t: $H_{t-1} \rightarrow S_2$. Mixed (behavior) strategies are σ_i^t: $H_{t-1} \rightarrow \Sigma_i$.

This class of games has been studied by Fudenberg-Kreps-Maskin (1988). We summarize their results here for the convenience of the reader.

[2] We have in mind a situation where the short-run players can learn about previous play without being physically present, either through word-of-mouth or perhaps from written sources. Note that if the short-run players could not condition their play on previous outcomes every equilibrium of the repeated game would be a sequence of static equilibria.

Let B: $\Sigma_1 \rightrightarrows \Sigma_2$ be the correspondence that maps mixed strategies by player one in the stage game to the best responses of player two. Because each short-run player cares only about the period-t outcome, in any equilibrium of $G(\delta)$ each period's play must lie in the graph of B.

Fudenberg-Kreps-Maskin prove that a kind of "folk theorem" obtains for games with a single long-run player. Specifically, let V_1 be the set of payoffs for player one attainable in the graph of B when player one is restricted to pure strategies. Let v_1 be player one's minimax value in the game in which moves by player two that are not best responses to some play by player one are deleted. Then any payoff in V_1 that gives player one more than v_1 can be attained in a sequential equilibrium if δ is near enough to one.

The point of this paper is to argue that if the game $G(\delta)$ is perturbed to allow for a small amount of incomplete information, then there is a far narrower range of Nash equilibrium payoffs. Roughly speaking, in the perturbed game the long-run player can exploit the possibility of building a reputation to pick out the equilibrium he likes best.

Specifically, define

$$g_1^* = \max_{s_1 \in S_1} \min_{\sigma_2 \in B(s_1)} g_1(s_1, \sigma_2) \tag{1}$$

and let s_1^* satisfy

$$\min_{\sigma_2 \in B(s_1^*)} g_1(s_1^*, \sigma_2) = g_1^*.$$

We call g_1^* player one's Stackelberg payoff and s_1^* a Stackelberg strategy. Note that there can be several Stackelberg strategies. These definitions differ slightly from the usual formulation, because when player two has several best responses we choose the best response that player one likes least. In the usual definition of Stackelberg equilibrium, the follower is assumed to play the best response that the leader most prefers. This is natural in many games with a continuum of strategies, because the leader can make the follower strictly prefer the desired response by making a small change in his strategy. In our finite setting, player one cannot break player two's indifference in this way, so to have a lower bound on player one's payoff, we need to allow for the case where player two is "spiteful" and chooses the best response that player one prefers least. On the other hand, in generic finite games there are no ties, and our definition coincides with the usual one.

In the next section we show that if there is positive prior probability that player one is a type that is certain to play s_1^* in every period, then player one's worst Nash equilibrium payoff is not much lower than g_1^* when δ is close to one. Because definition (1) limits player one to pure strategies, for some kinds of games the bound g_1^* is not useful. For example, in matching pennies, where the winning payoff is 1, $g_1^* = -1$, while each player can guarantee himself 0. Moreover, as shown by Fudenberg-Kreps-Maskin (1988), if we allowed mixed strategies in (1), not only could the resulting Stackelberg payoff be increased, it might actually exceed the highest payoff of any Nash equilibrium of the unperturbed game.

Our (1988) paper extends our results here to show that player one can in fact maintain a reputation for playing a mixed strategy if the corresponding "Stackelberg type" has positive probability.

3. THE PERTURBED GAME

This section introduces the perturbed game and gives the first version of our bounds on the long-run player's Nash equilibrium payoff. Section 4 gives examples to show that this bound cannot in general be improved on, and that there are generally many Nash equilibria.

In the perturbed game, player one knows his own payoff function, but the short-run players do not. We represent their uncertainty about player one's payoffs using Harsanyi's (1967) notion of a game of incomplete information. Player one's payoff is identified with his "type" $\omega \in \Omega$. It is common knowledge that the short-run players have (identical) prior beliefs about ω, represented by a probability measure μ on Ω. In this section we restrict attention to perturbed games with a countable number of types, so that $\Omega = \{\omega_0, \omega_1, \omega_2, \dots\}$ is a measure space in the obvious way. Section 7 allows for an uncountable number of types.

The period payoffs in the *perturbed game* are the same as in the unperturbed game, except that player one's period t payoff $g_1(s_1, s_2, \omega)$ may now depend on his type. A pure strategy for player one in the perturbed game is a sequence of maps $s_1^t\colon H_{t-1} \times \Omega \to S_1$, specifying his play as a function of the history and his type; a mixed (behavior) strategy is $\sigma_1^t\colon H_{t-1} \times \Omega \to \Sigma_1$. Otherwise the definition of the perturbed game is the same as that of the unperturbed game. The perturbed game is denoted $G(\delta, \mu)$ to emphasize its dependence on the long-run player's discount factor and on the beliefs of the short-run players.

Let type ω_0 have the preferences corresponding to the unperturbed game $G(\delta)$, so that if player one is type ω_0 then

$$g_1(s_1, s_2, \omega_0) = g_1(s_1, s_2).$$

In some cases it will be most natural for $\mu(\omega_0)$ to be near to unity; in others it will not. The magnitude of $\mu(\omega_0)$ is inessential for our argument, but we do require that it be strictly positive.

For any action $s_1 \in S_1$, let the event $\omega(s_1)$ be a type ω such that playing s_1 in every period is strictly dominant in the repeated game.[3] If repeated play of s_1 is strictly dominant in the repeated game, Nash equilibrium requires that (if $\omega(s_1)$ and h_t have positive probability) $s_1^{t+1}(h_t, \omega(s_1)) = s_1$. The event $\omega^* = \omega(s_1^*)$

[3] If s_1 was merely dominant in the stage game, it would not necessarily be dominant in the repeated game. In the prisoner's dilemma, for example, defection is dominant at each stage, but certainly is not a good strategy against a tit-for-tat opponent. To ensure that s_1 is dominant in the repeated game, let the stage-game payoffs for type $\omega(s_1)$ be:

$$g_1(s_1, s_2, \omega(s_1)) = g_1(s_1, \bar{s}_2, \omega(s_1)) > g_1(\bar{s}_1, \bar{s}_2, \omega(s_1)) \quad \text{for all} \quad \bar{s}_1 \neq s_1 \quad \text{and} \quad s_2, \bar{s}_2.$$

In other words, player one's payoff is independent of player two's action if he plays s_1, and is strictly more than he can get if he plays any other strategy.

means that player one strictly prefers to play the "Stackelberg strategy" s_1^*. We will say that this event corresponds to player one being the Stackelberg type.

Each (possibly mixed) strategy profile (σ_1^t, σ_2^t) induces a probability distribution π over $(S_1 \times S_2)^\infty \times \Omega$. Notice that this probability distribution contains many null events corresponding to deviations from the support of these strategies. Let h^* be the event $s_1^t = s_1^*$ for all t. Let π_t^* be the random variable $\pi(s_1^t = s_1^* | h_{t-1})$, that is, the probability attached by the short-run players to the Stackelberg strategy being played after h_{t-1}. Let $n(\pi_t^* \leq \bar{\pi})$ be the number (possibly infinite) of the random variables π_t^* for which $\pi_t^* \leq \bar{\pi}$. This is clearly a random variable.

The following lemma about statistical inference is at the heart of our results. It asserts that if s_1^* is always played, there is a fixed finite bound on the number of periods in which player two will believe s_1^* is unlikely to be played.

LEMMA 1: *Let* $0 \leq \bar{\pi} < 1$. *Suppose* $\mu(\omega^*) = \mu^* > 0$, *and that* (σ_1^t, σ_2^t) *are such that* $\pi(h^* | \omega^*) = 1$. *Then*

$$\pi\left[n(\pi_t^* \leq \bar{\pi}) > \log \mu^* / \log \bar{\pi} | h^*\right] = 0,$$

and for any infinite history h such that the truncated histories h_t *all have positive probability and such that* s_1^* *is always played,* $\pi(\omega^* | h_t)$ *is nondecreasing in t.*

REMARK: The lemma does not show that $\mu(\omega^* | h_t)$ converges to 1. In other words, the short-run players need not become convinced that the long-run player is type ω^*, even if the long-run player plays s_1^* in every period. Rather, the short-run players become convinced that the long-run player will play as if he were type ω^*.

PROOF:[4] Because the strategy spaces are finite and h^* has positive probability we may restrict attention to histories h_t that occur with positive probability, and such that player one has always played s_1^*. Fix such a history and for $1 \leq \tau < t$, let h_τ be the history h_t truncated at time τ.

Observe that by Bayes law

$$\pi(\omega^* | h_t) = \frac{\pi(\omega^* | h_{t-1}) \pi(h_t | \omega^*, h_{t-1})}{\pi(h_t | h_{t-1})}. \tag{2}$$

Moreover, since given h_{t-1} the choices of players one and two are independent,

$$\pi(h_t | h_{t-1}) = \pi\left(h_t^1, h_t^2 | h_{t-1}\right) = \pi\left(h_t^1 | h_{t-1}\right) \pi\left(h_t^2 | h_{t-1}\right). \tag{3}$$

Since player two's choice at t depends on player one's play only through h_{t-1}, and $\pi(\omega^*, h_{t-1}) > 0$,

$$\pi\left(h_t^2 | h_{t-1}\right) = \pi\left(h_t^2 | \omega^*, h_{t-1}\right). \tag{4}$$

[4] We are grateful to Paul Milgrom for suggesting a simplified exposition of this proof, which resulted as well in a better bound.

Also, since $\pi(h^* | \omega^*) = 1$, $\pi(h_t^1 | \omega^*, h_{t-1}) = 1$, so

(5) $$\pi\left(h_t^2 | \omega^*, h_{t-1}\right) = \pi\left(h_t | \omega^*, h_{t-1}\right).$$

Substituting (5) into (4) into (3) into (2) yields

(6) $$\pi(\omega^* | h_t) = \frac{\pi(\omega^* | h_{t-1})}{\pi\left(h_t^1 | h_{t-1}\right)}.$$

Since h_t^1 is distinguished from h_{t-1}^1 by the occurrence of the event $s_1^t = s_1^*$, $\pi(h_t^1 | h_{t-1}) = \pi(s_1^t = s_1^* | h_{t-1}) = \pi_t^*$. It follows that

(7) $$\pi(\omega^* | h_t) = \frac{\pi(\omega^* | h_{t-1})}{\pi_t^*}$$

where $\pi(\omega^* | h_{-1})$ is defined to be $\pi(\omega^*)$. Consequently, $\pi(\omega^* | h_t)$ is non-decreasing, and increases by a factor of at least $1/\bar{\pi}$ whenever $\pi_t^* \leq \bar{\pi}$. Since $\pi(\omega^*) = \mu^*$ is positive by assumption and $1 \geq \pi(\omega^* | h_t)$, we conclude that $n(\pi_t^* \leq \bar{\pi}) < k$, almost surely (given h_1^*), provided

(8) $$\mu^* / \bar{\pi}^k > 1.$$

Taking logs yields the conclusion of the lemma. *Q.E.D.*

Since the perturbed game has countably many types and periods, and finitely many actions per type and period, the set of Nash equilibrium pay-offs is a closed nonempty set. This follows from the standard results on the existence of mixed strategy equilibria in finite games, the assumption that the long-run player discounts future utilities, and the limiting results of Fudenberg and Levine (1983, 1986). Consequently, we may define $\underline{V}_1(\delta, \mu, \omega_0)$ to be the least payoff to a player one of type ω_0 in any Nash equilibrium of the perturbed game $G(\delta, \mu)$. (This minimum is taken over all mixed strategy equilibria.) Our main theorem says that if the long-run player is patient relative to the prior probability $\mu(\omega^*)$ that he is "tough," then he can achieve almost his Stackelberg payoff. Moreover, the lower bound on the long-run player's payoff is independent of the preferences of the other types in Ω to which μ assigns positive probability.

THEOREM 1: *Assume* $\mu(\omega_0) > 0$, *and that* $\mu(\omega^*) \equiv \mu^* > 0$. *Then there is a constant* $k(\mu^*)$ *otherwise independent of* (Ω, μ), *such that*

$$\underline{V}_1(\delta, \mu, \omega_0) \geq \delta^{k(\mu^*)} g_1^* + \left(1 - \delta^{k(\mu^*)}\right) \min g_1.$$

PROOF: Consider the strategy for player one of always playing s_1^*. We will show that player one can assure himself of at least the bound in the theorem, so he must get at least this much in any Nash equilibrium.

Since the reaction correspondence $B(\sigma_1)$ is upper hemicontinuous we know that each element of $B(\sigma_1^t)$ is near to an element of $B(s_1^*)$ when π_t^* is sufficiently near to one. Moreover, since player two has a finite number of pure strategies, if σ_2 is near to an element of $B(s_1^*)$ then it must place probability close to one on pure strategies that are contained in $B(s_1^*)$. Since player two must be indifferent between all strategies that he is willing to assign positive probability, we conclude there is a probability $\bar{\pi} < 1$ such that $B(\sigma_1^t)$ is contained in $B(s_1^*)$ whenever $\pi_t^* > \bar{\pi}$.

Set $k(\mu^*) = \log \mu^* / \log \bar{\pi}$. In any Nash equilibrium type ω^* will play s_1^* on the equilibrium path. It then follows from Lemma 1, that if type ω_0 plays s_1^* always, there are at most $k(\mu^*)$ occasions where player two plays outside of $B(s_1^*)$, corresponding to the events $\pi_t^* \leq \bar{\pi}$. In the worst case these events occur at the beginning of the game where they are discounted least. This gives the desired bound. *Q.E.D.*

The same proof holds immediately for finitely-repreated games, including the case where $\delta = 1$: If there are enough periods, player one's average payoff cannot be much below the Stackelberg level.

While we have assumed that player two's payoffs are common knowledge, a simple extension allows us to interpret each player two's choice of s_2 as a choice of a strategy mapping his privately-known type into an action. Under this interpretation, player one's Stackelberg strategy s_1^* is the one that maximizes player one's expected payoff, given that each type of player two chooses a short-run best response. In the chain-store game, for example, if there is a sufficiently high probability that player two is a type which will enter whether or not player one is expected to fight, then player one's Stackelberg action is to acquiesce. (See Milgrom-Roberts (1982) and Fudenberg-Kreps (1987).)

It is also unimportant that player one's payoff be common knowledge in the unperturbed game. Let player one's possible types in the unperturbed game be $\theta \in \Theta$, with $g_1 = g_1(s_1, s_2, \theta)$ and g_2 independent of θ. Now construct a perturbed game with type space Ω, and assume that Θ is contained in Ω. For each θ, let $s_1^*(\theta)$ be that type's Stackelberg action, that is, its most preferred action in the graph of B, and assume that *each* of the events $\omega(s_1^*(\theta))$ has positive prior probability. Theorem 1 shows that each θ can attain nearly its Stackelberg payoff $g_1^*(\theta)$ when δ is near to one.

For notational reasons, we assumed that types in the perturbed game have stationary payoffs independent of time and history. As can be seen from Lemma 1, this assumption is irrelevant. Indeed, even the ω^* type(s) can have nonstationary payoffs, provided that playing s_1^* forever is strictly dominant for the repeated game.

More strongly, even the unperturbed game need not be stationary. In a nonstationary game we may define the Stackelberg payoff to be the greatest average present value obtainable when the short-run player plays a static best response in every period. The nature of the best response may, however, be time dependent. In this case, the Stackelberg strategy may also be time-dependent, and the Stackelberg type is one who has the Stackelberg strategy as dominant in the repeated game. The argument is similar to that above, except that now the critical probability $\bar{\pi}$ will be indexed by time. Provided only that it is bounded uniformly away from one, that is, $\bar{\pi}_t \leq \bar{\pi} < 1$ for all t, Lemma 1 implies Theorem

1 remains valid. An application along these lines may be found in Example 2 below.

David Kreps has pointed out an alternative interpretation of our results for those who are uncomfortable with the assumption that the Stackelberg type has strictly positive probability. Instead, we could suppose that there is a fixed finite subset of Ω' of Ω. Each type $\omega' \in \Omega$ has some strategy s_1' as a dominant strategy in the repeated game. Then type ω_0 can compute its payoff from committing itself to play any s_1', and our proof shows that if δ is close to unity then type ω_0 can approximate the payoff to the s_1' it most prefers.

Along these lines, we note that it is not essential that the types ω' being imitated have dominant strategies s_1' in the repeated game. It is enough to prove that they play s_1' in every equilibrium. We can exploit this observation to show that our results follow from weaker assumptions on μ, combined with refinements of Nash equilibrium. If, for example, we are willing to restrict attention to sequential equilibria, it is sufficient to prove that the type ω' will play s_1' in every sequential equilibrium. It then follows if the probability of type ω' is fixed and greater than zero, any other type that is sufficiently patient can do almost as well in any sequential equilibrium as he could by publicly committing to s_1'.

One aspect of Theorem 1 that deserves emphasis is that given μ^*, the bound on the long-run player's payoff is independent of the prior probabilities of other types. This allows, for example, there to be types that try to "trick" the short-run player by playing Stackleberg for a few periods and then switching to another strategy. Since the bound has greatest significance as $\delta \to 1$, uniformity is of some importance. One interpretation of δ increasing towards one is that the time periods become very short. In this interpretation, it is not very natural to hold the probability distribution over types fixed: If some type plays a certain way for a year and then switches, the number of periods until it switches will increase as $\delta \to 1$. Theorem 1 shows that our bound obtains even if we adjust the relative probability of the non-Stackleberg types as $\delta \to 1$ to account for this effect.

Finally, it is natural to ask whether Theorem 1 extends to the case in which s_1^* is a mixed strategy by the long-run player. In Fudenberg-Levine (1988), we show that a similar bound does obtain. The major difference is that Lemma 1 does not hold. Since the short-run players do not know if the long-run players actually played s_1^*, the probability of ω^* is not nondecreasing. Instead, it can be shown to follow a special sort of supermartingale, and in place of Lemma 1 we can show that if s_1^* is always played, for every $\varepsilon > 0$, there exists a k so that the probability that player two fails to play a best response more than k times is no more than ε. Consequently, a very patient player can insure that he gets almost g_1^* with very high probability.

4. EXAMPLES

We now present some examples to illustrate the power and the limitations of our result. Example 1 explores the values of μ^* and $\bar{\pi}$ in the original chain store game. Example 2 uses some variants of the chain store game to illustrate the advantage of our technique of proof: We obtain asymptotic results without the need to explicitly solve for the sequential equilibria. Example 3 shows that the equilibria need not be unique.

EXAMPLE 1: To understand how our payoff bound depends on δ and μ^*, we examine a simple simultaneous move version of Selten's (1977) celebrated chain store paradox. In this game the short-run player may choose to be in or out; the long-run player to acquiesce or fight. If the short-run player stays out, he gets 0 and the long-run player gets 3; if the short-run player is in, and the long-run player acquiesces, the long-run player gets 1, the short-run player 2. If the short-run player is in, and the long-run player fights, the long-run player gets 0 and the short-run player -1 (the sequential move version of this game is in Figure 4).

Let minmax g_1 be the minmax value for the long-run player given that the short-run player is playing a best response to some mixed strategy by the long-run player. In this example, minmax $g_1 = 1$, and $g_1^* = 3$. Let $g_1(\delta)$ be the bound in Theorem 1. Then $(g_1(\delta) - \text{minmax}\, g_1)/(g_1^* - \text{minmax}\, g_1)$ is a natural measure of the tightness of our bound: It measures how much the long-run player is assured of gaining relative to the most he can gain from precommitment. Alternatively, we subtract this ratio from one to obtain a measure of player one's loss relative to g_1^*, obtaining

$$\frac{g_1^* - g_1(\delta)}{g_1^* - \text{minmax}\, g_1} = \frac{(1-\delta^k)(g_1^* - \min g_1)}{g_1^* - \text{minmax}\, g_1} \approx \frac{kr(g_1^* - \min g_1)}{g_1^* - \text{minmax}\, g_1}$$

where r is the interest rate, and $\delta^k \approx 1 - kr$ for δ near 1.

In the example $g_1^* = 3$, $\min g_1 = 0$ and minmax $g_1 = 1$, so the loss is $kr \cdot 3/2$. Moreover, if $\bar{\pi}$, the probability of a fight, exceeds 2/3, it is strictly best for the short-run player to stay out. Our bound k on the number of periods of entry is $\log \mu / \log \bar{\pi}$ so that $\mu^* = \bar{\pi}^k = (2/3)^k$. It is interesting to note that this bound, taken over all Nash equilibria of all finite or infinite horizon perturbed games with a prior μ^* of always fight, is the same as that obtained by Kreps-Wilson (1982) and Milgrom-Roberts (1982) for the finite-horizon sequential equilibria of the game with two types. Because our bound is logarithmic in μ^*, the μ^* required to guarantee no more than k periods of entry falls quite rapidly in k. If we wish the loss to be less than 12%, so $kr \leq .08$, and the game is played monthly with annual discount factor 12%, an 8 month wait means a loss of no more than 10%. On the other hand $(2/3)^8 = 0.04$, so a 4% chance of facing a Stackelberg long-run player means it will require at most 8 periods to establish a Stackelberg reputation.

EXAMPLE 2: Consider the following infinitely-repeated version of the chain-store game. Each period, a short-run entrant decides whether or not to enter and the long-run incumbent chooses whether to fight or to acquiesce. For conformity with the assumptions of Theorem 1, we assume that these choices are made

entrant size	gain to fighting instead of acquiesing
big	4p - 2
medium	3p - 1
small	(5p-1)/2

FIGURE 1

simultaneously, and that at the end of the period the incumbent's choice is revealed, whether or not entry in fact occurred. The entrants differ in two ways. First, each period's entrant is either "strong" or "weak." Strong entrants always enter, weak ones have payoffs described below. Each period's entrant is weak with probability p, independent of the others, and being strong or weak is private information. Second, each entrant is one of three "sizes," big, medium, or small. Each entrant has probability z_b of being big, z_m of being medium-sized, and z_s of being small, and the entrants' sizes are public information. To preserve a stationary structure, we imagine that the incumbent learns the period t entrant's size at the start of period t. Each weak entrant receives 1 if it stays out, 0 if there is a fight, and 2 if it enters and the incumbent acquiesces. Thus weak entrants enter if the probability that they will be fought is less than $1/2$. The incumbent receives 2 if it acquiesces and 4 if no entry occurs. The incumbent receives $3/2$ if it has to fight a small entrant, 1 if it has to fight a medium one, and 0 if it has to fight a large one. Thus in the unperturbed game, there is an equilibrium in which all entrants enter and the incumbent acquiesces to all entry. The previous papers on the chain-store game had only one size of entrant, and so the possible "reputations" the incumbent would want (that is, the possible Stackelberg actions) to establish are to always fight and to always acquiesce.

To find the Stackelberg strategy here, we compute in Figure 1 the difference in payoffs between fighting and acquiescing, given that the entrants play their best response. If, for example, $1/3 < p < 1/2$, the Stackelberg strategy is to fight the small and medium-sized entrants, and acquiesce to the large ones. If the only types of the incumbent are the original one ω_0, and a type ω_1 that will always fight, then this reputation need not be credible: The first time the incumbent concedes to a large entrant he reveals that he is not type ω_1. But if there is a nonzero prior probability that the incumbent is a type that acquiesces only to large entrants, then the incumbent can develop a reputation for playing in this way.

We can modify Example 2 in several ways. First, let us consider how our result extends to nonstationary environments. Imagine that there are only medium size firms and that $p = 0.01$ in odd-numbered periods, and $p = 0.60$ in even-numbered periods. Player one's *constant* Stackelberg strategy is then to always acquiesce, because the cost (-0.97) of fighting in the odd periods outweighs the potential gain (0.80) from entry deterrence in the even ones. However, if we allow player one the opportunity to maintain a reputation for nonstationary play, he can do much better by fighting in the even periods and acquiesing in the odd

	L	R
U	1, 1	0, 0
D	0, 0	10, 10

FIGURE 2

ones. If the prior distribution on Ω assigns positive probability to player one fighting in even periods and acquiescing in the odd ones, the proof of Theorem 1 extends in the obvious way to show that the long-run player can maintain the desired reputation.

EXAMPLE 3: Consider the 2×2 game in Figure 2. If $\mu(\omega = \omega_0)$ is near to one, then, regardless of δ, the game $G(\delta, \mu)$ has several sequential equilibria, all of which satisfy our bound. For concreteness, suppose that $0 < \varepsilon \leq 1/11$, and that μ puts weight $1 - \varepsilon$ on ω_0, weight ε on ω^* (the player who always plays D), and no weight on any other types. One equilibrium of this game has the short-run players playing R regardless of history, and both types of player one playing D regardless of history. Another equilibrium has the first period's player two playing L, and type ω_0 of player one playing U in the first period. Play from the second period on matches that in the first equilibrium. Neither player one nor the first period's player two has an incentive to deviate, as the first-period actions are a static Nash equilibrium, and not deviating gives player one his highest possible payoff from period two on. There is no point in player one imitating the ω^* type in the first period since he will do just as well beginning next period anyway.

Yet a third equilibrium has type ω_0 starting with U and reverting to D after period T, with player two switching to R at $(T+1)$ or whenever player one plays D. This is an equilibrium provided T is large enough and δ is small enough that $1 - \delta^T > 10\delta(1 - \delta^{T-1})$. In this case, the loss to player one of playing D in period one (of $1 - \delta$) exceeds the potential gain from convincing two that $\omega = s_1^*$ (of $10\delta - [(\delta - \delta^T) + 10\delta^T]$. Moreover, once player one has revealed in period one (by playing U) that he is not type ω^*, he cannot later build a reputation for being this type.

5. GENERAL DETERMINISTIC STAGE GAMES

We now consider nonsimultaneous move stage games. We start with an example showing the long-run player may do much less well than predicted by the simultaneous-move theory. This may seem surprising, because the chain-store game considered by Kreps-Wilson (1982) and Milgrom-Roberts (1982) has sequential moves and not simultaneous ones. We will see that their positive results are due to the particular nature of the payoffs in the chain-store game.

EXAMPLE 4: This example has the same extensive form as the sequential-move version of the chain store game, but different payoffs. Player two begins by choosing whether or not to purchase a good from player one. If he chooses not to

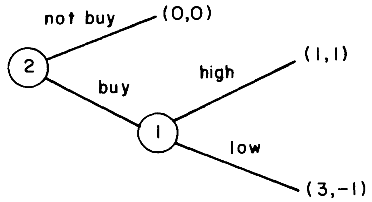

FIGURE 3.—A quality game.

buy, both players receive zero. If he buys, player one can produce high quality or low. High quality gives each player 1, while low quality gives player one a payoff of 3 and player two a payoff of -1 (see Figure 3). The Stackelberg outcome here is for player one to "promise" to choose high quality, so that all the customers will purchase. Thus if Theorem 1 extended to this game it would say that if there is a positive prior probability μ^* that player one is a type that always produces high quality, then if δ is near enough to unity, player one's payoff must be very close to 1 in any Nash equilibrium.

This extension is clearly false. Take $\mu^* = .01$ and $\mu(\omega_0) = .99$, and consider the following strategies. The high quality type ω^* produces high quality, and the "sane" type ω_0 is to produce low quality as long as no more than a single short-run player has made a purchase, and produces high quality beginning with the second time a short-run player purchases. The short-run players do not purchase unless a previous short-run player has, in which case they purchase as long as every previous purchaser except the first gets high quality. Otherwise they revert to not purchasing. Given the play of the long-run player, the short-run players are correct to not buy and so player one does not have the opportunity to demonstrate that he will produce high quality. Moreover, the first time a short-run player purchases, the long-run player will optimally produce low quality, because future players will purchase regardless of the quality that the first purchaser receives. While constructing a Nash equilibrium would be sufficient to show that Theorem 1 does not extend, it is interesting to note that these strategies form a sequential equilibrium when combined with the following beliefs $\mu(\omega^*|h^t)$. The beliefs are: $\mu(\omega^*|h^t) = .01$ if no purchase has been made, $\mu(\omega^*|h^t) = 0$ if low quality has ever been purchased, and $\mu(\omega^*|h^t) = 1$, otherwise. Facing these strategies a single long-run customer might be tempted to purchase once or twice to gather information, but myopic customers will not make this investment. Thus we see that for general stage games it is not true that player one can ensure almost his Stackelberg payoff. This poses a potential problem for the analysis of Kreps (1984), who uses a similar example to argue that firms may arise precisely to serve as a vehicle for this type of reputation.

Let us explain why the problem raised in Example 4 does not arise in the chain store paradox (Figure 4). There, the one action the entrant could take that would "hide" the incumbent's strategy is precisely the action corresponding to the Stackelberg outcome: Whenever the entrant does not play like a Stackelberg

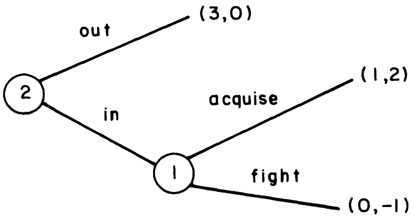

FIGURE 4.—The classical chain store paradox.

follower, the incumbent has a method of demonstrating that the entrant's play is "mistaken." The proof of Theorem 1 invoked Bayesian inference only in those periods where s_2^t does not belong to $B(s_1^*)$; whether the incumbent's strategy is revealed in periods where $s_2^t \in B(s_1^*)$ is irrelevant.

We have two responses to the problem posed by this example. The first is to follow Fudenberg-Maskin (1986) and Fudenberg-Kreps-Levine (1988) and examine perturbations with the property that every information set of the stage game is reached with positive probability. (See Fudenberg-Kreps-Levine for an explanation of the perturbations involved.) In this case by playing his Stackelberg action s_1^* in every period, player one ought to eventually force play to the Stackelberg outcome.

Our second response is to weaken Theorem 1. As we just saw, if the stage-game is not simultaneous-move, the long-run player may not have the opportunity to develop the reputation he would desire. 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 one's choice of normal-form strategy s_1. However when both players use pure strategies the information revealed about player one's play is deterministic. Let $O(s_1, s_2)$ be the subset of S_1 corresponding to strategies s_1' of player one such that (s_1', s_2) leads to the same terminal node as (s_1, s_2). We will say that these strategies are *observationally equivalent*. In the example, player one's Stackelberg action s_1^* was to produce high quality but, given that player two chooses not to risk dealing with him, player one had no way of establishing a reputation for Stackelberg play. The problem was that while "do not buy" was not a best response to s_1^*, "do not buy" is a best response to "low quality" and high and low quality are observationally equivalent when player two does not purchase.

This suggests the following generalization of Theorem 1: For any s_1, always playing s_1 should eventually force player two to play a strategy s_2 which is a best response to some s_1' in $O(s_1, s_2)$. That is, for each s_1 let $W(s_1)$ satisfy

$$W(s_1) = \{s_2 | \text{for some } \sigma_1' \text{ with support in } O(s_1, s_2), s_2 \in B(\sigma_1')\}. \tag{9}$$

In other words, $W(s_1)$ is the set of pure strategy best responses for player two to beliefs about player one's strategy that are consistent with the information revealed when that response is played. Then if δ is near to one, player one should be able to ensure approximately

$$g_1^* \equiv \max_{s_1} \min_{s_2 \in W(s_1)} g(s_1, s_2). \tag{10}$$

Notice that the fact that W is defined to contain only the pure strategy best responses by the short-run player does not effect the min $g_1(s_1, \sigma_2)$. As before, let s_1^* be a strategy that attains g_1^*, and let ω^* be the event that player one's best strategy in $G(\delta, \mu)$ is to always play s_1^*. Note that if the stage-game is simultaneous move, $W(s_1)$ consists of the pure strategy points in $B(s_1)$, and the definitions of g_1^* and s_1^* reduce to those we gave earlier.

THEOREM 2: *Let g be as described above. Assume $\mu(\omega_0) > 0$ and that $\mu(\omega^*) \equiv \mu^* > 0$. Then there is a constant $k(\mu^*)$, independent of μ, such that*

$$\underline{V}_1(\delta, \mu, \omega_0) \geq \delta^{k(\mu^*)} g_1^* + (1 - \delta^{k(\mu^*)}) \min g_1.$$

Before giving the proof, let us observe that this result, while not as strong as the assertion in Theorem 1 that player one can pick out his preferred payoff in the graph of B, does suffice to prove that player one can develop a reputation for "toughness" in the sequential-move version of the chain store game. Consider the extensive form in Figure 4 above. In this game $B(\text{fight}) = \{\text{out}\}$ and $B(\text{acquiesce}) = \{\text{in}\}$. Also, $O(\text{fight, out}) = O(\text{acquiesce, out}) = \{\text{acquiesce, fight}\}$, while $O(\text{fight, in}) = \{\text{fight}\}$ and $O(\text{acquiesce, in}) = \{\text{acquiesce}\}$.

First, we argue that $W(\text{fight}) = B(\text{fight})$. 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 two plays "in." Consequently, no strategy placing positive weight on "in" is in $W(\text{fight})$.

Finally, since player one's Stackelberg action with observable strategies is fight, and $W(\text{fight}) = B(\text{fight})$, the fact that only player one's realized actions, and not his strategy, is observable does not lower our bound on player one's payoff.

PROOF OF THEOREM 2: If $s_2 \notin W(s_1^*)$, then there is a $\bar{\pi}(s_2)$ such that s_2 is not a best response to any σ_1' which places weight at least $\bar{\pi}$ on $O(s_1^*, s_2)$. Let $\bar{\pi} \equiv \max \bar{\pi}(s_2)$. Each time player two plays an s_2' outside of $W(s_1^*)$, the observed outcome will be one that had prior probability less than $\bar{\pi}$. The proof then follows from Lemma 1. *Q.E.D.*

6. GENERAL FINITE STATE GAMES

This section extends the basic argument of Theorem 1 to the case with several interacting short-run players in each period.

The stage game is now a finite n-player simultaneous move game

$$g: S_1 \times S_2 \times \cdots \times S_n \to R^n,$$

with player one the only long-run player. Since each of the short-run players must play a short-run best response, in equilibrium each period's outcome must lie in the best response sets of all of the short-run players, that is, it must be a Nash equilibrium in the $(n-1)$ player game induced by fixing a (possibly mixed) strategy by player one. Consequently, we let $B\colon \Sigma_1 \rightrightarrows \Sigma_2 \times \cdots \times \Sigma_n$ be the Nash correspondence. This definition of $B(\sigma_1)$ agrees with our previous definition for the case of a single short-run player. The corresponding definition of g_1^* is

$$g_1^* = \max_{s_1} \min_{\sigma_{-1} \in B(s_1)} g_1(s_1, \sigma_{-1}). \tag{11}$$

As before, let s_1^* be a Stackelberg action for player one, that is, an action that attains g_1^*.

This situation is much the same as with a single short-run player, and as one would expect, the approach of Theorem 1 can be readily extended. There is only one minor complication: in the proof of Theorem 1 we argued that since the set $B(s_1^*)$ contained all the best responses to s_1^*, then there was a probability $\bar{\pi} < 1$ such that if player one was expected to play s_1^* with probability exceeding $\bar{\pi}$, player two would choose an action in $B(s_1^*)$. With several short-run players, the Nash correspondence $B(\cdot)$ need not be constant in the neighborhood of s_1^*, but it is still upper hemi-continuous. Thus, when player one plays s_1^* and his opponent plays a Nash equilibrium for some σ_1 near to s_1^*, the lowest player one's payoff can be is approximately g_1^*.

As before, let ω_0 be the type whose payoffs are as in the unperturbed game, and $\omega^* = \omega(s_1^*)$ be the event that player one has "s_1^* forever" as a dominant strategy.

THEOREM 3: *Fix a game $G(\delta)$ with several short-run players and consider a perturbed version $G(\delta, \mu)$. Assume $\mu(\omega_0) > 0$, and that $\mu(\omega^*) \equiv \mu^* > 0$. Then for all $\varepsilon > 0$, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$*

$$\underline{V}_1(\delta, \mu, \omega_0) \geq (1-\varepsilon) g_1^* + \varepsilon \min g_1.$$

Moreover, $\underline{\delta}$ depends on μ only through μ^*.

PROOF: For any $\pi \in (0,1]$ let $B(\pi, s_1^*)$ be the set of all Nash equilibria for the short-run players given any strategy for player one that puts probability at least π on s_1^*:

$$B(\pi, s_1^*) = \{\sigma_{-1} | \text{for some } \sigma_1 \text{ with } \sigma_1(s_1^*) \geq \pi, \sigma_{-1} \in B(\sigma_1)\} \tag{12}$$

and let

$$d(\pi) = \min_{\sigma_{-1} \in B(\pi, s_1^*)} g_1(s_1^*, \sigma_{-1}) \tag{13}$$

be the function that bounds how low player one's payoff can be when he plays s_1^* and the short-run players choose actions in $B(\pi, s_1^*)$. By definition $d(1) = g_1^*$. Moreover, $d(\pi)$ is clearly nondecreasing. We further claim that $d(\pi)$ is continuous in a neighborhood of 1. To see this, suppose to the contrary that there is an

$\varepsilon > 0$ and a sequence $\pi^n \to 1$ such that for all n, $d(\pi^n) < g_1^* - \varepsilon$. Then there is a sequence $\sigma_{-1}^n \in B(\pi^n, s_1^*)$ such that $g_1(s_1^*, \sigma_{-1}^n) < g_1^* - \varepsilon$. Extracting a convergent subsequence from the σ_{-1}^n, and using the upper hemicontinuity of $B(\cdot)$, we conclude that there is a $\bar{\sigma}_{-1} \in B(s_1^*)$ with $g_1(s_1^*, \bar{\sigma}_{-1}) < g_1^* - \varepsilon$, which contradicts the definition of g_1^*.

To prove the theorem, fix an $\varepsilon > 0$, and choose $\bar{\pi}$ such that $d(\bar{\pi}) > (1 - \varepsilon/2) g_1^* + (\varepsilon/2) \min g_1$. From Lemma 1 if player one chooses the strategy of always playing s_1^*, there is a bound $k(\mu^*, \bar{\pi})$, independent of δ, on how many times the short-run players can choose actions that are not in $B(\bar{\pi}, s_1^*)$. Then in all but $k(\mu^*, \bar{\pi})$ periods, the long-run player gets at least $(1 - \varepsilon/2) g_1^* + \varepsilon/2 \min g_1$. This yields a bound of

$$\delta^k g_1^* (1 - \varepsilon/2) + \delta^k (\varepsilon/2) \min g_1 + (1 - \delta^k) \min g_1.$$

The theorem now follows by choosing $\underline{\delta}$ close enough to one that $\underline{\delta}^k \geq (1 - \varepsilon)/(1 - \varepsilon/2)$. *Q.E.D.*

7. GAMES WITH A CONTINUUM OF STRATEGIES

We turn attention now to the case where players have a continuum of strategies in each period. Two complications arise in the analysis. First, it is no longer true that merely because the short-run player puts a large probability weight on the Stackelberg strategy, he must play a best response to it. However, he must play an ε-best response, and this is sufficient for our purposes. Second, it is not sensible to suppose that *a priori* the short-run player places positive weight on the Stackelberg strategy: instead we assume that all neighborhoods of the Stackelberg strategy have positive probability. Nash equilibrium, then, does not require the short-run player to optimize against the Stackelberg type, since that type occurs with probability zero. Instead, we must work with a sequence of types that converge to the Stackelberg type.

We return to the basic simultaneous move model. Our description of the perturbed and unperturbed game is unchanged with two exceptions. First, S_1 and S_2 are now assumed to be compact metric spaces, rather than finite sets, and Ω may be an arbitrary measure space. Second, the payoff maps g_1: $S_1 \times S_2 \times \Omega \to R$ and g_2: $S_1 \times S_2 \to R$ are assumed to be continuous on $S_1 \times S_2$. Next, the definitions of the Stackelberg payoffs and strategies must be changed. Because we are working with a continuum of types, it is no longer reasonable to assume that a single type ω_0 has strictly positive probability. Instead, we must consider types ω_0 belonging to a positive probability set Ω_0, with a different Stackelberg payoff $g_1^*(\omega_0)$ and Stackelberg strategy $s_1^*(\omega_0)$ for each ω_0. Thus we define for each ω_0

$$(14) \qquad g_1^*(\omega_0) = \sup_{s_1 \in S_1} \min_{\sigma_2 \in B(s_1)} g_1(s_1, \sigma_2, \omega_0).$$

Recall that in the previous sections we defined g_1^* as the maximum over a set of responses as opposed to the supremum. In the present setting with a continuum of actions, we must use the supremum because the function

$\min_{\sigma_2 \in B(s_1)} g(s_1, \sigma_2, \omega_0)$ need not be continuous. To see the problem, suppose that the short-run player has two strategies H and L, and that the long-run player chooses $s_1 \in [0,1]$. If the short-run player plays H, the long-run player gets $1 + s_1$; if the short-run player plays L, the long-run player gets s_1. On the other hand, the short-run player's payoff to H is s_1, and his payoff to L is $2s_1 - 1$. Consequently, if $s_1 < 1$, the unique best response for the short-run player is H; if $s_1 = 1$, he is indifferent between H and L. Thus $\min_{\sigma_2 \in B(s_1)} g_1(s_1, \sigma_2)$ is equal to $1 + s_1$ if $s_1 < 1$ but equal to 1 if $s_1 = 1$. However, we shall show that the supremum is obtainable in the limit.

With the definition of the type ω_0 Stackelberg payoff in hand, the definitions of the Stackelberg strategies is straightforward. Since S_1 is by assumption compact, we let $s_1^*(\omega_0)$ be a limit of a sequence s_1^m such that $\min_{\sigma_2 \in B(s_1^m)} g_1(s_1, \sigma_2, \omega_0) \to g_1^*(\omega_0)$. We also need to consider ε-best responses by player two. Define B_ε: $\Sigma_1 \rightrightarrows \Sigma_2$ to be the correspondence that maps mixed strategies by player one in the stage game g to ε-best responses of player two. That is, if $\sigma_2 \in B_\varepsilon(\sigma_1)$, then $g_2(\sigma_1, \sigma_2) + \varepsilon \geq g_2(\sigma_1, \sigma_2')$ for all $\sigma_2' \in \Sigma_2$. Let d denote the distance metric. We define

$$g_1^*(\varepsilon, s_1, \omega_0) = \min_{\substack{\sigma_2 \in B(s_1') \\ d(s_1', s_1) \leq \varepsilon}} g_1(s_1', \sigma_2, \omega). \tag{15}$$

In other words, we allow ε-best responses to strategies that differ from s_1 by up to ε. This function has one key property:

LEMMA 2: $\lim_{\varepsilon \to 0} g_1^*(\varepsilon, s_1, \omega_0) = g_1^*(0, s_1, \omega_0)$.

REMARK: Because of the discontinuity discussed above, $g_1^*(0, s_1^*, \omega_0)$ need not equal $g_1^*(\omega_0)$.

PROOF: If the lemma fails, there exist sequences $\varepsilon^m \to 0$, $s_1^m \to s_1$ and ε^m-best responses σ_2^m to s_1^m such that $\lim g_1(s_1^m, \sigma_2^m, \omega_0) < g_1^*(0, s_1, \omega_0)$. Since S_2 is compact, Σ_2 is weakly compact, and we may assume $\sigma_2^m \to \sigma_2$. Since g_1 is weakly continuous $\lim g_1(s_1^m, \sigma_2^m, \omega_0) = g_1(s_1, \sigma_2, \omega_0) < g_1^*(0, s_1, \omega_0)$. From the definition of $g_1^*(0, s_1, \omega_0)$ it is clear that σ_2 cannot be a best response to s_1; that is, there exists $\bar{\sigma}_2$ with $g_2(s_1, \bar{\sigma}_2) \geq g_2(s_1, \sigma_2) + \varepsilon$ for some $\varepsilon > 0$. However, since g_2 is weakly continuous, we have $g_2(s_1^m, \bar{\sigma}_2) \to g_2(s_1, \bar{\sigma}_2)$ and $g_2(s_1^m, \sigma_2^m) \to g_2(s_1, \sigma_2)$. This implies for large enough m

$$g_2(s_1^m, \bar{\sigma}_2) > g_2(s_1^m, \sigma_2^m) + \varepsilon/2 \geq g_2(s_1^m, \sigma_2^m) + \varepsilon^m,$$

contradicting the fact that σ_2^m is an ε^m-best response to s_1^m. *Q.E.D.*

To prove that as $\delta \to 1$, the long-run player can get nearly the Stackelberg payoff, we must describe our assumption on μ, the distribution over types. Recall that $\omega(s_1)$ is the type that has s_1 as a dominant strategy. As we remarked at the beginning of this section, we do not wish to assume that the Stackelberg type corresponding to ω_0 has positive probability. Instead we will assume that there is positive probability near $\omega(s_1^*(\omega_0))$. More precisely, let $\mu^*(d', s_1)$ denote the probability that μ assigns to the types $\omega(s_1')$ corresponding to the strategies s_1' that are within distance d' of s_1. We say that a subset Ω_0 of Ω has *positive Stackelberg probability* if for each $\omega_0 \in \Omega_0$ there exists a $d'(\omega_0) > 0$ such that for all s_1 with $d(s_1, s_1^*(\omega_0)) \leq d'(\omega_0)$ and all $d'' > 0$, $\mu^*(d'', s_1) > 0$. This condition will be satisfied if μ has a positive density over all types of the form $\omega(s_1)$.

THEOREM 4: *If Ω_0 has positive Stackelberg probability, then for almost all $\omega_0 \in \Omega_0$ and all $\varepsilon > 0$ there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$*

$$\underline{V}_1(\delta, \mu, \omega_0) \geq (1-\varepsilon) g_1^*(\omega_0) + \varepsilon \min g_1(\omega_0).$$

Moreover, $\underline{\delta}$ depends on μ only through $\mu(\cdot)$.

PROOF: The reason that the theorem holds only for almost all $\omega_0 \in \Omega_0$ is that type ω_0's equilibrium payoff $\underline{V}_1(\delta, \mu, \omega_0)$, a conditional expectation, is only defined for almost all ω_0. Fix an ω_0 for which $\underline{V}_1$ is defined, and choose $\bar{s}_1$ so that $|g_1^*(0, \bar{s}_1, \omega_0) - g_1^*(\omega_0)| < \varepsilon(g_1^*(\omega_0) - \min g_1(\omega_0))/3$ and $d(\bar{s}_1, s_1^*(\omega_0)) \leq d'(\omega_0)$. This is possible by the definition of $s_1^*(\omega_0)$. Choosing a small number ε' notice that the probability of the set of s_1 with $d(s_1, \bar{s}_1) < \varepsilon'$ is positive by assumption. It follows that for some s_1 with $d(s_1, \bar{s}_1) < \varepsilon'$, player two must respond optimally following the history h_t that results when s_1 is played repeatedly. In particular, since of the types $\omega(s_1')$ with $d(s_1', s_1) < \varepsilon'$, it is almost surely true in equilibrium that only type $\omega(s_1)$ will actually play s_1, we can choose s_1 so that after observing s_1 played, the type $\omega(s_1)$ will have positive conditional probability. It is clear that Lemma 1 continues to hold with infinite action spaces. Consequently, we conclude that for every $\bar{\pi} < 1$ there is a k such that if player one always plays $\bar{s}_1$, there is probability zero that $\pi(s_t^1 = \bar{s}_1 | h_{t-1}) \leq \bar{\pi}$ more than k times. Next, for each α, there is some $\bar{\pi} < 1$ such that player two will play an action in $B_\alpha(\bar{s}_1)$ whenever $\pi(s_1^\tau = \bar{s}_1 | h_{\tau-1}) > \bar{\pi}$. It follows that

$$\underline{V}_1(\delta, \mu, \omega_0) \geq \delta^k g_1^*(\alpha, \bar{s}_1, \omega_0) + (1-\delta^k) \min g_1(\omega_0)$$

where k depends only on α. By Lemma 2, we may choose α so that

$$|g_1^*(\alpha, \bar{s}_1, \omega_0) - g_1^*(0, \bar{s}_1, \omega_0)| < \varepsilon(g_1^*(\omega_0) - \min g_1(\omega_0))/3.$$

We conclude that for some k,

$$\delta^k g_1^*(\omega_0)(1 - 2\varepsilon/3) + \delta^k(2\varepsilon/3) \min g_1(\omega_0) + (1-\delta^k) \min g_1(\omega_0).$$

The theorem now follows by choosing $\underline{\delta}$ close enough to 1 that $\underline{\delta}^k \geq (1-\varepsilon)/(1-2\varepsilon/3)$. *Q.E.D.*

Let us conclude by observing that Theorem 4 can be extended along the lines of Section 5 to cover general stage games. This simply involves replacing the set $B_\varepsilon(s_1)$ with the set $W_\varepsilon(s_1)$ of strategies of player two that are a ε-best response to beliefs that are consistent with the information revealed when that response is played.

EXAMPLE 5: The observation that general stage games with a continuum of actions may be treated this way enables us to examine the sequential play of the Kreps-Wilson (1982) two-sided predation game, which was analyzed in Fudenberg-Kreps (1987). The stage game is played on the interval [0, 1], with each player choosing a time to concede if the other player is still fighting. If the short-run players are unlikely to be "tough," the Stackelberg action for the long-run player is to commit to fight until the finish ($t = 1$), which induces the "weak" short-run players to concede immediately. This game is not a simultaneous move: If the first player two concedes immediately, the others will not learn how long player one would have been willing to fight. However, as in the simple predation game, this does not pose additional complications. Whenever a short-run player does not play as a Stackelberg follower, the long-run player will have a chance to demonstrate that he is "tough." We conclude that if the long-run player is patient he can do almost as well as if he could commit himself never to concede.

Department of Economics, MIT, Cambridge, MA 02139, U.S.A.
and
Department of Economics, UCLA, Los Angeles, CA 90024, U.S.A.

Manuscript received July, 1987; final revision received November, 1988.

REFERENCES

AUMANN, R., AND S. SORIN (1987): "Cooperation and Bounded Rationality," forthcoming in *Games and Economic Behavior*.

FUDENBERG, D., AND D. KREPS (1987): "Reputation and Simultaneous Opponents," *Review of Economic Studies*, 54, 541–568.

FUDENBERG, D., D. KREPS, AND D. LEVINE (1988): "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, 354–380.

FUDENBERG, D., D. KREPS, AND E. MASKIN (1988): "Repeated Games With Long-Run and Short-Run Players," MIT Working Paper #474.

FUDENBERG, D., AND D. LEVINE (1983): "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory*, 31, 251–268.

——— (1986): "Limit Games and Limit Equilibria," *Journal of Economic Theory*, 38, 261–279.

——— (1988): "Reputation, Unobservable Strategies, and Active Supermartingales," MIT Working Paper #484.

FUDENBERG, D., AND E. MASKIN (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533–554.

HARSANYI, J. (1967): "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159–182, 320–334.

KREPS, D. (1984): "Corporate Culture and Economic Theory," Mimeo, Stanford Graduate School of Business.

KREPS, D., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245–252, 486–502.

KREPS, D., AND R. WILSON (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

MILGROM, P., AND J. ROBERTS (1982): "Predation, Reputation and Entry Deterence," *Econometrica*, 50, 443–460.

SELTEN, R. (1977): "The Chain-Store Paradox," *Theory and Decision*, 9, 127–259.