Review of Economic Studies (1992) 59, 561-579
0034-6527/92/00280561$02.00


# Maintaining a Reputation when Strategies are Imperfectly Observed

DREW FUDENBERG
*Massachusetts Institute of Technology*

and

DAVID K. LEVINE
*University of California, Los Angeles*

*First version received November* 1988; *final version accepted May* 1991 (*Eds.*)

This paper studies reputation effects in games with a single long-run player whose choice of stage-game strategy is imperfectly observed by his opponents. We obtain lower and upper bounds on the long-run player's payoff in any Nash equilibrium of the game. If the long-run player's stage-game strategy is statistically identified by the observed outcomes, then for generic payoffs the upper and lower bounds both converge, as the discount factor tends to 1, to the long-run player's Stackelberg payoff, which is the most he could obtain by publicly committing himself to any strategy.

## 1. INTRODUCTION

We consider "reputation effects" in a game in which a single long-run player faces a sequence of short-run opponents, each of whom plays only once but is informed of the outcomes of play in previous periods. A number of papers have studied such models since the idea was introduced by Kreps-Wilson (1982) and Milgrom-Roberts (1982) in their studies of the chain-store game. Put briefly, the idea is that if a player continually plays the same action, his opponents will come to expect him to play that action in the future. Moreover, if the opponents are myopic, as for example the short-run entrants in the chain-store game, then once they become convinced that the long-run player is playing a fixed stage-game strategy they will play a best response to that strategy in subsequent periods. Foreseeing this response, the long-run player may choose to "invest in his reputation" by playing the strategy even when doing so incurs short-run costs, provided the costs are outweighed by the long-run benefit of influencing his opponents' play.

Intuitively, one might expect that the benefits of investing in reputation will outweigh the costs when the long-run player is sufficiently patient. This intuition is most clear if the long-run player's stage-game strategy is perfectly observed by his opponents, for then the long-run player's investments have a direct and predictable effect on the evoluation of his reputation. In many situations of interest, though, the long-run player's stage-game strategy is imperfectly observed. This is the case if the long-run player uses a mixed stage-game strategy, if he is subject to moral hazard, or if his stage-game strategy prescribes actions for contingencies that need not arise when the stage game is played. In all of these cases the short-run players must try to infer the long-run player's past play from the observed information, so that the link between the long-run player's choices and the

The second improvement over our earlier paper is that we now allow for the possibility that the outcome of play is only statistically determined by the long-run player's action, so that the long-run player is subject to "moral hazard" in trying to maintain his reputation. One example of this is the model of Cukierman–Meltzer (1986), where a long-run central bank is trying to maintain a reputation for restraint in the control of the money supply. Individuals do not observe the bank's action, which is the rate of money growth, but instead observe realized inflation, which is influenced by money growth and also by a stochastic and unobserved shock. Thus unexpectedly high inflation could either mean that the shock was high or that the bank increased the money supply more rapidly than had been expected, which might seem to make it more difficult for reputation effects to emerge. Our results show that the addition of moral hazard does not change the basic reputation-effects intuition in that the limiting value of our payoff bounds is independent of the amount of noise in the system, so long as the oucome permits the "statistical identification" of the long-run player's play. (However, the noise can lower the long-run player's equilibrium payoff for a fixed value of the discount factor.)

Note that the generalizations to mixed-strategy reputations and to games with moral hazard are quite similar in a formal sense, as in both cases the complication is that the observed outcome reveals only imperfect information about the long-run player's unobserved strategy: in the case where actions are observed, as in the chain-store game, the long-run player's realized action will not reveal the randomizing probabilities that the long-run player used. This is why it is natural to consider the two generalizations in the same paper.

A third way this paper improves on our earlier work is that we now obtain an upper bound on the long-run player's payoff in addition to the lower bound. The upper bound converges, as the long-run player's discount factor approaches one, to the long-run player's "generalized Stackelberg payoff", which is a generalization of the idea of the Stackelberg payoff. The generalized Stackelberg payoff can be greater than the limit of the lower bound of the long-run player's equilibrium payoff, and in general games reputation effects do not always lead to sharp predictions. However, if the stage-game has simultaneous moves (or on the weaker condition of statistical identification), and the prior distribution has full support on the set of all commitment types (including those corresponding to mixed strategies) then for the generic payoffs the upper and lower bounds *both* converge to the Stackelberg payoff, and reputation effects have very strong implications indeed. This conclusion emphasizes the difference between games such as the chain-store example, with a single long-run player, and games with several long-run players, such as the repeated prisoner's dilemma considered by Kreps *et al.* (1982): With several long-run players the limit set of equilibrium payoffs with reputation effects need not be a singleton, and can depend on the relative probabilities of various "commitment types" (Aumann–Sorin (1989), Fudenberg–Maskin (1986)).

Here is an intuition for our results for games in which the realized stage-game strategies are observed, that is, simultaneous-move games without moral hazard. Fix a Nash equilibrium $\sigma^*$ of the game, and suppose that the long-run player (of whatever type) decides to play the equilibrium strategy $\sigma_1^*(\bar{\omega})$ of a type $\bar{\omega}$ in the support of the prior distribution. Since the short-run players are myopic, they will play a best response to $\sigma_1^*(\bar{\omega})$ in any period where they expect player 1's stage-game strategy to be close to $\sigma_1^*(\bar{\omega})$. Conversely, if the short-run players do not play a best response to $\sigma_1^*(\bar{\omega})$, then one would expect them to be "surprised" if that strategy is indeed played. That is, we would expect them to increase the probability they assign to the long-run player being type $\bar{\omega}$. This is reflected in the fact that when $\sigma_1^*(\bar{\omega})$ is a pure strategy, the posterior probability that $\omega = \bar{\omega}$ increases in any period where the forecast differs from $\sigma_1^*(\bar{\omega})$. If $\sigma_1^*(\bar{\omega})$ is a mixed strategy (or, more generally, the long-run player's stage-game strategy is not directly observed) the analogous statement is that there is a non-negligible probability that the outcome causes the short-run players to revise their beliefs by a non-negligible amount. More precisely, the martingale convergence theorem implies that for any $\varepsilon$ there is a $K(\varepsilon)$ such that with probability $(1-\varepsilon)$ the short-run players will expect player 1 to play $\sigma_1^*(\bar{\omega})$ in all but $K(\varepsilon)$ periods.

To obtain the desired payoff bounds, we must strengthen this assertion by finding an upper bound on the $K(\varepsilon)$ that holds uniformly over all Nash equilibria for a given discount factor and also over all discount factors. Then, when the long run player is very patient, what his opponents play in the fixed number $K(\varepsilon)$ of "bad" periods is unimportant.

Finally, we get an upper bound on payoffs by taking $\bar{\omega}$ to be the long-run player's true type, while we get a lower bound by taking $\bar{\omega}$ to be a "type" committed to playing the Stackleberg strategy.

## 2. THE MODEL

The long-run player, player 1, plays a fixed stage game against an infinite sequence of different short-run player 2's. In the stage game player 1 selects an action $a_1$ from a finite set $A_1$, while that period's player 2 selects from a finite set $A_2$. Denote action profiles by $a \in A \equiv A_1 \times A_2$. The stage game is not required to be a simultaneous move, but is allowed to correspond to an arbitrary game tree, so that the "actions" should be thought of as contingent plans or pure strategies for the stage game. At the end of each period, the players observe a stochastic outcome $y$ which is drawn from finite set $Y$ according to the probability distribution $\rho(\cdot \mid a)$. This outcome is defined to include all of the information players receive about each others' actions. The case where actions are directly observed is modelled by identifying a distinct outcome $y(a)$ with each action profile, then setting $\rho(y(a) \mid a) = 1$.

There are two reasons that the outcome $y$ need not reveal the action profile. First, if the profile represents a strategy in an extensive-form stage game, then even if the outcome $y$ is deterministic it will not reveal how players would have played at information sets that were not reached under $a$. This possibility is illustrated in Figure 1, where the outcomes are identified with the terminal nodes of the stage game. Here the outcome "no sale" does not reveal the quality that would have been sold if the consumer had bought.

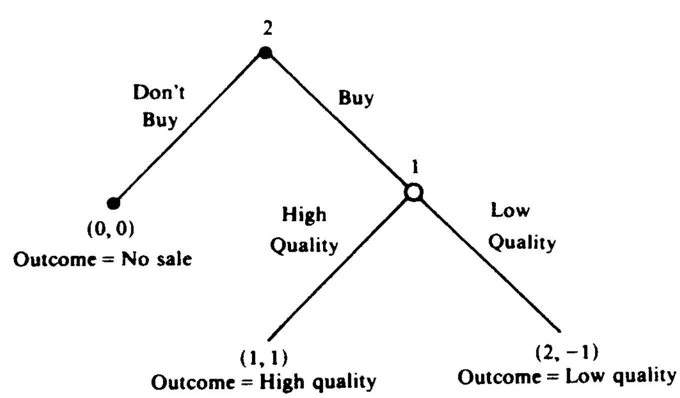


FIGURE 1

Second, even if the actions are uncontingent choices, the distribution of outcomes for a fixed action profile may be stochastic, so that the outcome gives only imperfect statistical information about the actions. This is the case for the example in the Cukierman-Meltzer (1986) paper on inflation and monetary policy.

Corresponding to the $A_i$ are the spaces $\mathscr{A}_i$ of mixed actions; when the mixed action profile is $\alpha \in \mathscr{A}_1 \times \mathscr{A}_2$ the resulting probability of $y$ is

$$\rho(y \mid \alpha) = \sum_{a \in A} \rho(y \mid a)\alpha_1(a_1)\alpha_2(a_2).$$

(Note that this formulation includes the special case where $A$ and $Y$ are isomorphic.)

The outcome $y$ has been defined to contain all of the information the short-run players receive about the long-run player's choice of action $a_1$. Accordingly, we require that their payoff depend on $a_1$ only through its influence on the distribution of $y$.[2] The short-run players all have the same expected utility function $u_2: Y \times A_2 \to \mathbf{R}$. Let

$$v_2(\alpha) = \sum_{a \in A, y \in Y} u_2(y, a_2)\rho(y \mid a)\alpha_1(a_1)\alpha_2(a_2)$$

denote the expected payoff corresponding to the mixed action $\alpha$. Each period's short-run player acts to maximize that period's expected payoff.

All players know the short-run players' payoff function. On the other hand, the long-run player knows his own payoff function, but the short-run players do not. We represent their uncertainty using Harsanyi's (1967) notion of a game of incomplete information. The long-run player's payoff is identified with his "type" $\omega \in \Omega$, where $\Omega$ is a metric space. It is common knowledge that the short-run players have (identical) prior beliefs $\mu$ about $\omega$, represented by a probability measure on $\Omega$.[3] As with short-run players, we suppose that the per-period payoff of the long-run player depends on the action $a_2$ of his opponent only through its influence on the distribution of $y$. We allow this utility $u_1(a_1, y, \omega, t)$ to be non-stationary, and assume that is bounded uniformly, so that for some $\underline{u} < \bar{u}$, $\underline{u} \leqq u_1(a_1, y, \omega, t) \leqq \bar{u}$ for all $\omega$ and $t$. The overall utility is the expected average discounted value

$$E(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u_1(a_1(t), y(t), \omega, t)$$

where $0 \leqq \delta < 1$. The normalization by $(1-\delta)$ places per-period and repeated game payoffs on the same scale. As in the case of the short-run player, we may define the expected payoff to a mixed action:

$$v_1(\alpha, \omega, t) = \sum_{a \in A, y \in Y} u_1(a_1, y, \omega, t)\rho(y \mid a)\alpha_1(a_1)\alpha_2(a_2).$$

Both long-run and short-run players can observe and condition their play at time $t$ on the entire past history of the realized outcomes. The long-run player can also condition his play on his private information and on his own past actions. Let $H_t$ denote the set of possible public histories (of outcomes) through time $t$, including the null history $h_0$. A pure strategy for the period-$t$ player 2 is a map $s_2^t: H_{t-1} \to A_2$, while the set of all such strategies is denoted $S_2^t$. Let $H_t^1$ denote the set of player 1's possible private histories (the past realizations of $a_1$) through time $t$. A pure strategy $s_1$ for any type $\omega$ of player 1 is a sequence of maps $s_1^t: H_{t-1} \times H_{t-1}^1 \to A_1$, specifying his play as a function of history; the set of all such $s_1$ is denoted $S_1$.

2. An alternative interpretation is that the outcome $y$ is defined to include the short-run player's payoff as well as a "signal" of the long-run player's action.

3. Throughout the paper all of our measure spaces are topological spaces endowed with the Borel $\sigma$-algebra.

Let $\Sigma_1$ and $\Sigma_2^t$ be the sets of probability distributions over $S_1$ and $S_2^t$, let $\Sigma_2 \equiv \times_{t=1}^{\infty} \Sigma_2^t$. Each $\sigma \in \Sigma_1 \times \Sigma_2$ gives rise to a probability distribution over sequences of actions and outcomes. Consequently we let $E_\sigma$ denote the expectation with respect to this distribution, and define

$$U_1(\sigma, \omega) \equiv E_\sigma(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u_1(a_1(t), y(t), \omega(t)$$

to be the expected utility to player 1.

Strategies for the long-run player must also account for the fact that there are many types. Since we wish to allow $\Omega$ to be infinite, we follow Milgrom-Weber (1985) in considering *distributional strategies* $\mathscr{s}_1 \in \mathscr{S}_1$. These are joint probability distributions over $\Omega \times S_1$ with the property that the marginal on $\Omega$ equals $\mu$. Each $\mathscr{s}_1$ and $\Omega' \subset \Omega$ with $\mu(\Omega') > 0$ gives rise to conditional strategies $\mathscr{s}_1(\Omega') \in \Sigma_1$ through integration: If $S_1' \subseteq S_1$, $\mathscr{S}_1(\Omega')(S_1') = \mathscr{s}_1(\Omega' \times S_1')/\mu(\Omega')$.

A ***Nash equilibrium*** can now be defined as a pair $\mathscr{s}_1, \sigma_2 \in \mathscr{S}_2 \times \Sigma_2$ so that $\sigma_2^t$ is a best response to $\mathscr{s}_1(\Omega)$ and so that if $(\omega, s_1) \in \text{support } \mathscr{s}_1$, then $s_1$ is a best response to $\sigma_2$ by type $\omega$.

Each $\sigma \in \Sigma_1 \times \Sigma_2$ induces a probability distribution over $H$, the public histories of infinite length. Moreover, each $h_t \in H_t$ may be identified with the subset of $h \in H$ that coincide with $t$ through time $t$. In this way, we may view the $H_t$ (which are finite sets) as sub-sigma algebras of the Borel sets in $H$, and view random variables on $H$ as stochastic processes on $(\{H_t\}_{t=1}^{\infty}, H)$. We shall adopt this point of view frequently in the sequel.

Since our main concern will be the evolution of the posterior probabilities over the long-run player's type, it is convenient to have a special notation to express the likelihood function for the event that $\omega$ lies in various subsets of $\Omega$. Fix $\mathscr{s}_1, \sigma_2$ and a subset $\Omega^+ \subset \Omega$, with $\mu(\Omega^+) > 0$. Let $\Omega^- = \Omega \backslash \Omega^+$ be the complement of $\Omega^+$. We let $\sigma_1 = \mathscr{s}_1(\Omega)$, $\sigma_1^+ = \mathscr{s}_1(\Omega^+)$ and $\sigma_1^- = \mathscr{s}_1(\Omega^-)$ be the induced probability distribution over strategies corresponding to all types, types in $\Omega^+$ and types in $\Omega^-$ respectively. We also set $\sigma = \sigma_1 \times \sigma_2$, $\sigma^+ = \sigma_1^+ \times \sigma_2$, $\sigma^- = \sigma_1^- \times \sigma_2$. The corresponding probability distributions on outcomes at time $t$ conditional on $h_{t-1}$ are denoted $p(h_{t-1})$, $p^+(h_{t-1})$ and $p^-(h_{t-1})$ respectively. These are families of random vectors on $H$. Similarly, we can set $\alpha_1(h_{t-1})$, $\alpha_1^+(h_{t-1})$, and $\alpha_1^-(h_{t-1})$ to be the time-$t$ probabilities of actions by player 1 given $h_{t-1}$, and set $\alpha_2(h_{t-1})$ to be the conditional probability of actions by player 2. Note that $\alpha_2(h_{t-1}) \equiv \sigma_2^t(h_{t-1})$ is independent of $\sigma_1$. Since the expected conditional expectation equals the expectation, we may calculate

$$U_1(\sigma^+, \omega) = E_{\sigma^+} \sum_{t=1}^{\infty} \delta^{t-1} v_1(\alpha_1^+(h_{t-1}), \sigma_2(h_{t-1}), \omega, t).$$

Finally, the conditional probability of a type in $\Omega^+$ under $\sigma$ given $h_{t-1}$ is denoted $\mu(\Omega^+ | h_{t-1})$.

The power of reputation effects depends on which reputations are *a priori* feasible and this depends on which types have positive prior probability. To model reputations for always playing a particular pure action, we use "commitment types" who prefer to always play that action. There are several ways of constructing the model so that the long-run player has the option of trying to maintain a reputation for playing a mixed action. The simplest way to do this supposes there are "commitment types" who like to play specific mixed actions. While this may not be completely implausible, it has the awkward feature that such types cannot be expected utility maximizers. Alternatively, reputations for mixed actions can be modelled in an expected utility framework with the following technical device. Let $\Omega_1 \subset \Omega$ be the "irrational types". We identify $\Omega_1$ with $\mathscr{A}_1 \times A_1^{\infty}$, the product of mixed actions with the space of sequences of actions. (The latter is a compact space in the product topology.) If $\omega = (\alpha_1, \{a_1(t)\}_{t=1}^{\infty})$ then $u_1(a_1, y, \omega, t) = \bar{u}$

if $a_1 = a_1(t)$ and $u_1(a_1, y, \omega, t) = \underline{u}$ if $a_1 \neq a_1(t)$, so that almost surely type $\omega$ will play $\{a_1(t)\}_{t=1}^{\infty}$ in a Nash equilibrium of the repeated game. Note that irrational types are expected utility maximizers, but have non-stationary time-additive preference: each strictly prefers a particular sequence of actions to all others. We let $\Omega(\alpha_1)$ be the set $\{\omega \in \Omega_1 \mid \omega = (\alpha_1, \{a_1(t)\}_{t=1}^{\infty})$ for some $\{a_1(t)\}_{t=1}^{\infty}\}$. Given $\mu$, probabilities conditional on the sets $\Omega(\alpha_1)$ exist for almost all $\alpha_1$. On the other hand, playing $\alpha_1$ independently in each period, by Kolmogorov's Theorem induces a unique probability distribution on $\Omega(\alpha_1)$. We assume that $\mu$ is such that the conditional probabilities on $\Omega(\alpha_1)$ are equal to this distribution almost surely. In this sense $\Omega(\alpha_1)$ can be viewed as a kind of "commitment type": conditional on $\omega$ lying in this set, the unique dominant strategy leads to a probability distribution over actions for the long-run player that is "as if" he mixed independently following $\alpha_1$.

The prior $\mu$ induces a measure $\eta$ on the set of mixed actions by $\eta(\mathscr{A}_1') = \mu(\{\bigcup_{\alpha_1 \in \mathscr{A}_1'} \Omega(\alpha_1)\})$. If the part of $\eta$ that is absolutely continuous with respect to Lebesgue measure is non-zero and has a density that is uniformly bounded away from 0, we say that *commitment types have full support.*

As an example, to model a single type who likes to randomize with $\alpha_1$ equals $(\frac{1}{2}, -\frac{1}{2})$ between actions $H$ and $T$, we introduce a set of types $\Omega(\alpha_1)$ corresponding to sequences $(H, H, H, \ldots)(T, H, H, \ldots)$, $(H, T, H, T, \ldots)$ and so forth, together with the induced probability distribution from i.i.d. coin flips.

## 3. SELF-CONFIRMING RESPONSES AND EQUILIBRIUM PAYOFFS

This section develops a theorem on the upper and lower bound of the long-run player's Nash equilibrium payoffs. The proof uses a result about Bayesian inference, proved in the next section, that provides uniform bounds on how often the short-run players can be "substantially wrong" in their forecast of the long-run players play.

Fix $\sigma_1$, $\sigma_2$ and $\Omega^+$ with $\mu(\Omega^+) > 0$. Note that $p(h_{t-1})$ is the forecast the period-$t$ player 2, makes about period-$t$ outcomes, knowing $\sigma_1 \times \sigma_2$. By way of contrast $p^+(h_{t-1})$ is what he would forecast if he knew that the long-run player's type is in $\Omega^+$. If $p$ is a probability distribution over outcomes, $m = 1, \ldots, M$, define $\|p\| \equiv \max_m \|p_m\|$. In the next section we prove the following result.

**Theorem 4.1.** *For every $\varepsilon > 0$, $\Delta_0 > 0$ and $\Omega^+$ with $\mu(\Omega^+) > 0$ there is a K depending only on these three numbers such that for any $\sigma_1$ and $\sigma_2$, under the probability distribution generated by $\sigma_1(\Omega^+)$, there is probability less than $\varepsilon$ that there are more than K periods with*

$$\|p^+(h_{t-1}) - p(h_{t-1})\| > \Delta_0.$$

Loosely speaking, if $\Omega^+$ is true, the short-run players forecast the outcome $y$ about as well in almost every period as they would if they knew that $\Omega^+$ was true. (This is loose because $p$ and $p^+$ depend on the short-run player's action $a_2$).

This section uses Theorem 4.1 to characterize the long-run player's equilibrium payoffs. To begin, we define what it means for the short-run player's action to be a best response to approximately correct beliefs about the distribution over *outcomes*, as opposed to beliefs about the long-run player's *action*. Because we will assume that commitment types have full support, all mixed actions by player 1 have positive probability in any Nash equilibrium. Thus no player 2 will ever choose a weakly dominated strategy, and we exclude these strategies in our definition of a best response.

*Definition.* A mixed action $\alpha_2$ is an *$\varepsilon$-confirmed best response to $\alpha_1$* if
(1) $\alpha_2$ is not weakly dominated[4]
(2) there exists $\alpha_1'$ such that
(a) $\alpha_2$ solves $\max_{\alpha_2'} v_2(\alpha_1', \alpha_2')$, and
(b) $\|\rho(\cdot|(\alpha_1, \alpha_2)) - \rho(\cdot|(\alpha_1', \alpha_2))\| < \varepsilon$.

We let $B_\varepsilon(\alpha_1)$ denote the set of all $\varepsilon$-confirmed best responses to $\alpha_1$. Note that $B_0(\alpha_1)$ is not the same as the set of all (undominated) best responses to $\alpha_1$, as there may be distinct strategies $\alpha_1$ and $\alpha_1'$ with $\rho(\cdot|(\alpha_1, \alpha_2)) = \rho(\cdot|(\alpha_1', \alpha_2))$, and $B_0(\alpha_1)$ then contains best responses to both $\alpha_1$ and $\alpha_1'$. For example, in the game of Figure 1, "buy" is the unique best response to "high quality", but "don't buy" is also in $B_0$ (high quality) as "don't buy" is a best response to "low quality", and the profiles (high quality, don't buy) and (low quality, don't buy) lead to the same terminal nodes. In the terminology of our (1989) paper the elements of $B_0(\alpha_1)$ are *generalized best responses.*

We now relate this to Nash equilibrium payoffs. First note that a Nash equilibrium exists in each finite-horizon truncation of the game (see Milgrom-Weber (1985)). Compactness and the fact that preferences are uniformly continuous in the product topology then implies the existence of a convergent sequence of truncated equilibria whose limit is an equilibrium in the infinite game. (See Fudenberg-Levine (1983), for example.) For a long-run player of type $\omega$, if $\mu(\{\omega\}) > 0$, we can define $\underline{N}_1(\delta, \omega)$ and $\bar{N}_1(\delta, \omega)$ to be the infinum and supremum of his payoff in any Nash equilibrium.

In equilibrium, if the commitment types have full support, the fact that short-run players play myopically implies that $\alpha_2(h_{t-1}) \in B_0(\alpha_1(h_{t-1}))$. Moreover, if

$$\| p^+(h_{t-1}) - p(h_{t-1})\| \leqq \Delta_0,$$

as in the conclusion of Theorem 4.1, then in equilibrium $\alpha_2(h_{t-1}) \in B_{\Delta_0}(\alpha_1^+(h_{t-1}))$.

We now focus our attention on a particular class of long-run players. Type $\omega_0$ is *time-stationary* if $u_1(a_1, y, \omega_0, t) = u_1(a_1, y, \omega, t')$ for all $t, t'$. For such a type $v_1$ is time-stationary as well. Our main theorem provides bounds on the equilibrium payoff of time-stationary types in terms of the "$\varepsilon$-least" and "$\varepsilon$-greatest" commitment payoffs, which we will now define. The *$\varepsilon$-least commitment payoff* is

$$\underline{v}_1(\omega, \varepsilon) \equiv \sup_{\alpha_1 \in \mathcal{A}_1} \inf_{\alpha_2 \in B_\varepsilon(\alpha_1)} v_1(\alpha_1, \alpha_2, \omega) - \varepsilon.$$

This is $\varepsilon$ less than the least that type $\omega$ gets by committing to any fixed strategy when the short-run player plays an $\varepsilon$-confirmed response. Note that the definition allows player 2 to choose the response player 1 likes least whenever he is indifferent between two or more responses. This is a pessimistic measure of the power of commitment. The *$\varepsilon$-greatest commitment payoff* is

$$\bar{v}_1(\omega, \varepsilon) = \sup_{\alpha_1 \in \mathcal{A}_1} \sup_{\alpha_2 \in B_\varepsilon(\alpha_1)} v_1(\alpha_1, \alpha_2, \omega).$$

Obviously

$$\bar{v}_1(\omega, \varepsilon) \geqq \underline{v}_1(\omega, \varepsilon).$$

For $\varepsilon = 0$ we call $\bar{v}_1(\omega, 0)$ the *generalized Stackelberg payoff.* Since the supremum is taken over all generalized best responses to $\alpha_1$, instead of only the best responses, the generalized Stackelberg payoff is at least as large as the usual Stackelberg payoff (modulo our restriction to undominated responses).

4. An action $\alpha_2$ is weakly dominated if there exists $\alpha_2'$ such that $v_2(\alpha_1, \alpha_2') \geqq v_2(\alpha_1, \alpha_2)$ for all $\alpha_1 \in \mathcal{A}_1$, with strict inequality for at least one $\alpha_1$.

7

If the observed outcomes correspond to the terminal nodes of an extensive-form stage game, then the generalized Stackelberg payoff is the same as the usual one. To see this, recall that if $\alpha_2$ is a generalized best response to $\alpha_1$, then there is an $\alpha_1'$ such that $(\alpha_1', \alpha_2)$ and $(\alpha_1, \alpha_2)$ generate the same probability distribution over outcomes. Thus, if outcomes correspond to terminal nodes, $(\alpha_1', \alpha_2)$ and $(\alpha_1, \alpha_2)$ generate the same distribution over terminal nodes and hence give player 1 the same expected payoff. Thus player 1 can do as well by playing $\alpha_1'$ and having player 2 play a best response to player 1's true action.

However, when the outcomes do not correspond to the terminal nodes, the generalized Stackelberg payoff can be strictly greater than the usual one.[5] As an example, consider a simultaneous-move game in which player 1 chooses $U$ or $D$, and player 2 chooses $L$ or $R$. There are three outcomes that occur in a deterministic manner: If $(U, L)$ or $(D, L)$ is played the outcome is $y_1$, if $(U, R)$ is played the outcome is $y_2$, while $(D, R)$ leads to $y_3$. Player 1's payoff function is $u_1(U, y_1) = 2$; $u_1(D, y_1) = 0$; $u_1(\cdot, y_2) = 0$; $u_1(\cdot, y_3) = 1$, player 2's payoffs are $u_2(L, \cdot) = 1$; $u_2(\cdot, y_2) = 2$ and $u_2(\cdot, y_3) = 0$. The strategic form is shown in Figure 2.

Player 1

| Player 2 | L | R |
|---|---|---|
| U | 2, 1 | 0, 2 |
| D | 0, 1 | 1, 0 |

Payoffs

| | L | R |
|---|---|---|
| U | $y_1$ | $y_2$ |
| D | $y_1$ | $y_3$ |

Outcome

FIGURE 2

Here the generalized Stackelberg payoff is 2, which is attained by player 1 playing $U$, and player 2 playing $L$. $L$ is a generalized best response to $U$, as $L$ is a best response to $D$, and $\rho(\cdot \mid (U, L)) = \rho(\cdot \mid (D, L))$. In contrast, the Stackelberg payoff is only 1, which is attained by $\alpha_1 = (\frac{1}{2}U, \frac{1}{2}D)$. We now state our main result.

**Theorem 3.1.** *If $\omega_0$ is a stationary type with $\mu(\{\omega_0\}) > 0$, and commitment types have full support, then for all $\varepsilon > 0$ there exists $K$ so that, for all $\delta$,*

$$(1-\varepsilon)\delta^K \underline{v}_1(\omega_0, \varepsilon) + [1-(1-\varepsilon)\delta^K]\underline{u} \leqq \underline{N}_1(\delta, \omega_0) \leqq \bar{N}_1(\delta, \omega_0)$$
$$\leqq (1-\varepsilon)\delta^K \bar{v}_1(\omega_0, \varepsilon) + [1-(1-\varepsilon)\delta^K]\bar{u}.$$

*Proof.* As we remarked earlier, because commitment types have full support, player 2 never plays a strategy that is weakly dominated by any mixed strategy. Consequently, in any Nash equilibrium $(\sigma_1, \sigma_2)$, if

$$\|p^+(h_{t-1}) - p(h_{t-1})\| \leqq \Delta_0, \qquad \alpha_2(h_{t-1}) \in B_{\Delta_0}(\alpha_1^+(h_{t-1})).$$

5. We thank an anonymous referee for pointing this out to us.

To establish the upper bound, take $\Omega^+=\{\omega_0\}$. We have

$$U_1(\sigma^+,\omega_0)=E_{\sigma^+}(1-\delta)\textstyle\sum_{t=1}^{\infty}\delta^{t-1}v_1(\alpha_1^+(h_{t-1}),\alpha_2(h_{t-1}),\omega_0)$$

as the payoff to $\omega_0$. Choosing $\Delta_0=\varepsilon$, we conclude from Theorem 4.1 that there is a $K$ such that with probability $(1-\varepsilon)$, player 2's equilibrium action $\alpha_2(h_{t-1})$ lies in $B_\varepsilon(\alpha_1^+(h_{t-1}))$ in all but $K$ periods. Since $\alpha_1^+(h_{t-1})$ is the expected play of type $\omega_0$ given history $h_{t-1}$ (averaging over the different private histories $h_{t-1}^1$ consistent with $h_{t-1}$), we conclude that type $\omega_0$'s expected equilibrium payoff is bounded by $\bar{v}_1(\omega_0,\varepsilon)$ in all but the $K$ "exceptional" periods. Since payoffs in the exceptional periods are bounded above by $\bar{u}$, and the present value is maximized if the payoffs $\bar{u}$ occur in the first $K$ periods, the upper bound follows.

To establish the lower bound choose $\varepsilon'>0$ so that if $|\alpha_1'-\alpha_1|\leqq\varepsilon'$ then $\|v_1(\alpha_1,\alpha_2,\omega_0)-v_1(\alpha_1',\alpha_2,\omega_0)\|<\varepsilon$ and $\|\rho(\cdot|(\alpha_1',\alpha_2))-\rho(\cdot|(\alpha_1,\alpha_2))\|\leqq\varepsilon/2$ for all $\alpha_2$. Such an $\varepsilon'$ exists since $v$ and $\rho$ are continuous functions on compact sets.) Fix an $\alpha_1$, and take $\Omega^+$ to be the union of $\Omega(\alpha_1')$ over $|\alpha_1'-\alpha_1|\leqq\varepsilon'$. Note that since $\alpha_1^+(h_{t-1})$ is a convex combination of $\alpha_1'$ satisfying $|\alpha_1'-\alpha_1|\leqq\varepsilon'$, $\|\alpha_1^+(h_{t-1})-\alpha\|\leqq\varepsilon'$. Note also that since the commitment types have full support, $\mu(\Omega^+)$ is bounded below by a $\underline{\mu}>0$ that is independent of $\alpha_1$ (but in general depending on $\varepsilon'$). Consequently we may find a $K$ as in the conclusion of Theorem 4.1 that is independent of $\alpha_1$.

Suppose that type $\omega_0$ adopts the strategy $\alpha_1^+$ associated with $\Omega^+$. Choosing $\Delta_0=\varepsilon/2$ we see that under $\sigma^+$ there is probability at least $(1-\varepsilon)$ that in all but $K$ exceptional periods $\alpha_2(h_{t-1})\in B_{\varepsilon/2}(\alpha_1^+(h_{t-1}))$. Since $\|\alpha_1^+(h_{t-1})-\alpha_1\|\leqq\varepsilon'$, $\|v_1(\alpha_1^+(h_{t-1}),\alpha_2,\omega_0)-v_1(\alpha_1,\alpha_2,\omega_0)\|<\varepsilon$, and so except in the exceptional periods, $\omega_0$ gets at least $\min_{\alpha_2\in B_\varepsilon(\alpha_1)}v_1(\alpha_1,\alpha_2,\omega_0)-\varepsilon$. The lower bound now follows from taking the supremum over $\alpha_1\in\mathscr{A}_1$. ‖

It might be thought that the upper bound is too weak, and in particular that when actions are observed, the highest equilibrium payoff for any discount factor should be the Stackelberg payoff $\bar{v}_1(\omega,0)$. However, this is not the case. For a fixed discount factor a type may receive "information rents" from the possibility of other types that give it a payoff higher than the Stackelberg level. Consider for example the matrix game in Figure 3, where player 1 chooses rows, player 2 chooses columns, and the matrix gives the payoffs of type $\omega_0$ of player 1 and player 2.

Here type $\omega_0$'s Stackelberg payoff is 1, and he would like to commit himself to $U$. Type $\omega_0$ would not like to commit himself to play $D$, because that strategy is strictly dominated. However, type $\omega_0$ would like player 2 to believe that he was playing $D$, for this induces the response of $L$, and allows him a payoff of 2. If the prior distribution places high probability on player 1 being a type who usually plays $D$, then player 2 will play 2 at least in the first few periods. Thus if $\delta$ is small, type $\omega_0$ could obtain a payoff of almost 2 by playing $U$ while player 2 played $L$. However, to obtain this payoff 1 must

| | | Player 1: L | Player 1: R |
|---|---|---|---|
| Player 2 | $U$ | 2, 0 | 1, 1 |
| | $D$ | −1, 1 | −1, 0 |

FIGURE 3

"fool" player 2. The intuition for the upper bound is that the short-run players cannot be fooled infinitely often.

One example from the literature where the long-run player does do better than his Stackelberg payoff for small $\delta$ is Benabou-Laroque (1988). They consider a model of an informed "insider" who knows if the state is "good" or "bad", and who can send a possibly dishonest report of his information to the market. The "sane" insider would like to mislead the market, but this cannot occur in equilibrium when his type is known; the sane type's commitment payoff corresponds to revealing no information. However, if the market believes there is positive probability that the insider will always report honestly, the sane type can play a mixed strategy that announces "good" when the state is "good" with probability less than $\frac{1}{2}$, but such that marginal distribution of reports, averaged over the insider's type, is that an announcement of "good" means the probability that the state is good exceeds $\frac{1}{2}$. Hence the market price will rise when the insider announces "good", and the "sane" insider can earn a rent by misleading the market. Benabou-Laroque are concerned with the nature of the equilibrium strategies for a fixed $\delta$ and not in the sorts of payoff bounds that we develop. However, it is interesting to note that our theorem implies that the sane type's payoff converges to the no communication payoff as $\delta$ converges to 1, regardless of the prior probability that the speculator is honest.

We now formalize the idea information rents should vanish in the limit as $\delta \to 1$. Let $\underline{N}_1(\omega)$ be the lim inf of $\underline{N}_1(\delta, \omega)$ and let $\bar{N}_1(\omega)$ be the lim sup of $\bar{N}_1(\delta, \omega)$.

**Corollary 3.2.** *If $\omega_0$ is a stationary type with $\mu(\{\omega_o\}) > 0$ and the commitment types have full support,*

$$\underline{v}_1(\omega_0, 0) \leqq \underline{N}_1(\omega_0) \leqq \bar{N}_1(\omega_0) \leqq \bar{v}_1(\omega_0, 0).$$

*Proof.* Letting $\delta \to 1$ and then $\varepsilon \to 0$ in Theorem 3.1 shows that we need only establish that $\liminf \underline{v}_1(\omega_0, \varepsilon) \geqq \underline{v}_1(\omega_0, 0)$ and $\limsup \bar{v}_1(\omega_0, \varepsilon) \leqq \bar{v}_1(\omega_0, \varepsilon)$. This in turn follows from the upper hemi-continuity of $B_\varepsilon(\alpha_1)$ in $\varepsilon$; that is $\varepsilon^n \to 0$, $\alpha_2^n \in B_{\varepsilon^n}(\alpha_1)$ implies that any limit point $\alpha_2$ of $\alpha_2^n$ lies in $B_0(\alpha_1)$. ||

When is the bound tight; that is, when are the lower and upper bounds equal? Roughly speaking, two conditions are required: the short-run players should care about how the long-run player plays, and the outcome $y$ must reveal "enough" statistical information about the long-run player's action.

To see how the $\underline{v}_1$ and $\bar{v}_1$ can differ if the short-run player does not care about the actions of the long-run player, consider a game in which $A_1$ is a singleton and $A_2 = \{0, 1\}$. The short-run player's payoff is zero no matter what he plays, while $u_1(a_1, a_2, \omega_0) = a_2$. Clearly $\underline{N}_1(\omega_0) = 0$ and $\bar{N}_1(\omega_0) = 1$ regardless of the discount factor or the presence of commitment types.

For the two bounds to be equal we need to exclude this type of game. The games we exclude are degenerate in the sense that they are non-generic in the space of payoff functions. The game is *non-degenerate* if there is no undominated pure action $a_2 \in A_2$ such that for some $\alpha_2 \neq a_2$,

$$v(\cdot, a_2) = v(\cdot, \alpha_2).$$

This condition rules out the degenerate example above, but is satisfied for an open dense set of payoffs: if $a_2$ is undominated and $v_2(\cdot, a_2) = v_2(\cdot, \alpha_2)$ with $a_2 \neq \alpha_2$, then in the game with $u_2^\varepsilon(y, a_2') = u_2(y, a_2')$ for $a_2' \neq a_2$, and $u_2^\varepsilon(y, a_2) = u_2(y, a_2) - \varepsilon$, $a_2$ is weakly dominated by $\alpha_2$. (It can also be shown that non-degenerate games have full measure in the space of payoff functions.)

Next, to see what can happen when the outcome does not reveal sufficient information about the long-run player's action, consider the following quality choice game adopted from Fudenberg-Levine (1989) (see also Figure 1). The long-run player chooses $a_1 \in$ {high quality, low quality}, and the short-run player may play $a_2 \in$ {do not buy, buy}. The possible outcomes are $Y =$ {no sale, buy high quality, buy low quality}, corresponding to the 3 terminal nodes of the extensive form. If the short-run player buys, the outcome is buy high quality or buy low quality according to $a_1$. If he does not buy, the outcome is no sale regardless of $a_1$. The short-run player gets 1 if he buys high quality, $-1$ if he buys low quality and 0 if no sale. Consider a type $\omega$ of long-run player who gets 1 if he sells high quality, 2 if low and 0 if no sale. Clearly $\bar{v}_1(\omega, 0) = 1{\cdot}5$, for if $\omega$ mixes $\frac{1}{2}-\frac{1}{2}$ between high and low, the short-run player is still willing to buy. If $\mu(\omega) \geqq 0{\cdot}5$, however, then it is a Nash equilibrium for the long-run player to play low quality, and for the short-run player to not buy, implying $\underline{v}(\omega, 0) = 0$ (which is the individually rational payoff). The long-run player cannot build a reputation for producing high quality because the short-run player never buys and so never observes the long-run player's action.

To rule out this possibility, we use the following condition: the game is *identified* if for all $\alpha_2$ that are not weakly dominated, $\rho(\cdot\,|\,\alpha_1, \alpha_2) = \rho(\cdot\,|\,\alpha_1', \alpha_2)$ implies $\alpha_1 = \alpha_1'$. This condition requires that distinct actions of the long-run player induce distinct distributions over outcomes. Clearly the game is identified if the long-run player's actions are observed, so that for every $y$ there is a unique $a_1$ independent of $a_2$ such that $\rho(y\,|\,a_1, a_2) > 0$. This is the case if the stage game is a one-shot simultaneous-move game, and there is no moral hazard.

Even if there is moral hazard, the game may be identified. Let $R(\alpha_1)$ be the matrix with columns corresponding to outcomes $y$, rows corresponding to actions $a_1$ for the long-run player, and entries $R_{a_1, y} = \rho(y\,|\,a_1, a_2)$. Since $\rho(\cdot\,|\,\alpha_1, \alpha_2) = \alpha_1 R(\alpha_2)$, if $R(\alpha_2)$ has full row rank for all undominated $\alpha_2$, the game is identified. It might seem that if player 1 has no more actions than there are outcomes, this condition is generically true in simultaneous move games. However, this is somewhat deceptive. While it is true that if the number of outcomes is at least the number of actions by the long-run player $R(\alpha_2)$ will generically have full row rank for all pure strategies, $R(\alpha_2)$ will generically have full row rank only for almost all mixed strategies.

As an example, consider a game in which both players have two actions, heads $H$ and tails $T$, and there are two outcomes, also called $H$ and $T$. If $\alpha_2 = H$, $y = a_1$. If $a_2 = T$, then the outcome is the opposite of whatever player 1 chose, that is, $(H, T)$ produces outcome $T$, while $(T, T)$ produces outcome $H$. Thus $R(\alpha_2)$ has full rank except for $\alpha_2 = (\frac{1}{2}H, \frac{1}{2}T)$: in this case each outcome has probability $\frac{1}{2}$ regardless of how the long-run player plays, so the game is not identified. Moreover, perturbing the information structure slightly will not make the game identified. Fortunately, many simultaneous-move economic games have information structures satisfying natural monotonicity assumptions that rule out this type of singularity.

A more interesting economic example is that of a repeated signalling game. Here the long-run player (worker) draws a type ("high productivity," "low productivity") $\tau_1$ from a finite set each period in an i.i.d. manner. He then makes a decision $d_1$ ("go to school", "do not go to school") based on his type. An action $a_1$ is a map from types to decisions $d_1 = a_1(\tau_1)$. The short-run player (firm) moves second and observes the decision of the long-run player, but not his type. He then makes a decision $d_2$. An action $a_2$ for the short-run player is a map from long-run player decisions to short-run player decisions

11

$d_2 = a_2(d_1)$. At the end of the period, after the short-run player's decision is final, the current type of long-run player $\tau_1$ is revealed to the current and all subsequent short-run players. Clearly, the game is identified, since any mixture over maps from types to decisions induces a unique distribution over pairs $(\tau_1, d_1)$, both observed *ex post.* Our result below implies that in the repeated signalling game, the worker can do as well as by committing to any map from type to schooling.

In contrast, if the long-run player moves after the short-run player and has more than one information set, the game will typically fail to be identified. Even if players observe the terminal node, so the short-run player observes the way the long-run player played, this will not reveal the stage-game strategy he chose. This is reflected in a non-generic $R(\alpha_2)$ matrix. In the game in Figure 1, if the short-run player plays "do not buy", the only possible outcome is "no sale", and the corresponding $2\times 3$ matrix $R(\alpha_2)$ has rank one.

On the other hand, we have assumed that the long-run player has many types, and that his type is private information. It may be reasonable to suppose that the same is true of the short-run players. If these types are also chosen independently from period to period, and there are sufficient variety of types, every sequence of moves of the short-run player will have positive probability. In this case, if the terminal node is observed, it is clear that the game is identified, although for some $\alpha_2$, $R(\alpha_2)$ may not have full row rank, and indeed, may not even have more rows than columns.

**Theorem 3.3.** *In a non-degenerate game that is identified, for any stationary type* $\omega_0$, $\underline{v}_1(\omega_0, 0) = \bar{v}_1(\omega_0, 0)$.

*Proof.* Recall that $B_0(\alpha_1)$ is the set of weakly undominated $\alpha_2$ such that there exists an $\alpha_1'$ with $\rho(\cdot\,|\,\alpha_1', \alpha_2) = \rho(\cdot\,|\,\alpha_1, \alpha_2)$ and such that $\alpha_2$ is a best-response to $\alpha_1'$. Since the game is identified the only such $\alpha_1'$ is $\alpha_1$, so $B_0(\alpha_1)$ is simply the undominated best responses to $\alpha_1$.

Therefore, it suffices to show that for $\alpha_2 \in B_0(\alpha_1)$ there exists a sequence $\alpha_1^n \to \alpha_1$ such that $B_0(\alpha_1^n) = \{a_2\}$. Now $a_2$ is by definition undominated and, by the hypothesis of non-degeneracy, does not yield the same vector of payoffs to player 2 as any other mixed strategy. Thus there exists an $\alpha_1'$ such that $a_2$ is a strict best response to $\alpha_1'$. Then, however, it is a strict best response to $\lambda\alpha_1 + (1-\lambda)\alpha_1'$ for all $0 \leqq \lambda < 1$. Let $0 < \lambda^n < 1$ with $\lambda^n \to 1$. Then $\alpha_1^n = \lambda^n\alpha_1 + (1-\lambda^n)\alpha_1' \to \alpha_1$, and $a_2$ is the unique best response to $\alpha_1^n$. ||

## 4. BAYESIAN INFERENCE AND ACTIVE SUPERMARTINGALES

We now demonstrate Theorem 4.1, stated in the previous section, that it is unlikely that forecasts of $y$ are wrong in many periods. We do so via several lemmas analyzing the odds ratio $[1-\mu(\Omega^+\,|\,h_t)]/\mu(\Omega^+\,|\,h_t)$ for arbitrary sets $\Omega^+$. If this odds ratio is low, $\Omega^+$ is likely to be true, so conditional forecasts of $y$ under $\sigma^+$ are close to those under $\sigma$. On the other hand, when conditional forecasts of $y$ are different under $\sigma^+$ than under $\sigma$, the odds ratio has a good chance of falling substantially if the true strategy is $\sigma^+$. Because the odds ratio is a supermartingale, it converges almost surely. Moreover, it is well known that the odds ratio sampled only at periods where the forecasts of $y$ are different under $\sigma$ and $\sigma^+$ is a supermartingale that converges to zero (see, for example, Neveu (1975)). We strengthen this observation to show that for a fixed difference between the conditional forecasts $p^+(h_{t-1})$ and $p(h_{t-1})$, the odds ratio converges to zero at a uniform rate, independent of the particular distributions $p^+$ and $p$.

We begin by defining families of scalar random variables $(p_t^+(h), p_t^-(h))$. Set $p_t^+ = p_m^+(h_{t-1})$ and $p_t^- = p_m^-(h_{t-1})$ if $y(t)$ is the m-th element of $Y$. (Recall that $h_t \in H_t$ is the finite history that coincides with $h$ through and including time $t$.) Define another family of random variables $L_t(h)$ as follows:

$$L_0(h) \equiv \frac{1-\mu(\Omega^+)}{\mu(\Omega^+)}.$$

$$L_t(h) \equiv \frac{p_t^-(h)}{p_t^+(h)} L_{t-1}(h).$$

It is well known that $L_t(h) = [1-\mu(\Omega^+|h_t)]/\mu(\Omega^+|h_t)$, which is the posterior odds ratio at the end of period $t$ under $\sigma^+$ that player 1 is not in $\Omega^+$. It is also well known that this odds ratios is a supermartingale under the distribution $\sigma^+$. We give a proof for completeness.

**Lemma 4.1.** *$L_t(h) = [1-\mu(\Omega^+|h_t)]/\mu(\Omega^+|h_t)$ and $(L_t, H_t)$ is a supermartingale under $\sigma^+$.*

*Proof.* The first claim is by definition true for $L_0$. Imagine it is true for $L_{t-1}$, then

$$[1-\mu(\Omega^+|h_t)]/\mu(\Omega^+|h_t) = p_t^-[1-\mu(\Omega^+|h_{t-1})]/[p_t^+(\Omega^+|h_{t-1})]$$
$$= (p_t^-/p_t^+)L_{t-1} = L_t.$$

To see that $L_t(h)$ is a supermartingale, recall that $p^+(h_{t-1})$ is the conditional probability over $y$ from $\sigma^+$ so that

$$E[L_t|L_{t-1}, h_{t-1}] = L_{t-1}\textstyle\sum_{m|p_m^+(h_{t-1})>0}[p_m^-(h_{t-1})/p_m^+(h_{t-1})]$$
$$= L_{t-1}\textstyle\sum_{m|p_m^+(h_{t-1})>0} \leqq L_{t-1}. \quad \|$$

Let $\Delta(h_{t-1}) = \|p^+(h_{t-1}) - p^-(h_{t-1})\|$ be the distance between the distributions over outcomes corresponding to $\Omega^+$ and $\Omega^- = \Omega\backslash\Omega^+$. Note that $\|p^+(h_{t-1}) - p(h_{t-1})\| \leqq \Delta(h_{t-1})$ since $p(h_{t-1})$ is a convex combination of $p^+(h_{t-1})$ and $p^-(h_{t-1})$.

Next we show that the odds ratio $L_t$ is likely to fall substantially when $\Delta(h_{t-1}) > \Delta_0$.

**Lemma 4.2.** *If $h_{t-1}$ has positive probability and $\Delta(h_{t-1}) > \Delta_0$ then under $\sigma^+$, $\Pr[(L_t(h)/L_{t-1}(h)) - 1 \leqq -\Delta_0/M | h_{t-1}] \geqq \Delta_0/M$.*

*Proof.* Note first that $L_t/L_{t-1} = p_t^-/p_t^+$ which is $p_1^-(h_{t-1})/p_1^+(h_{t-1})$ with probability $p_1^+(h_{t-1})$; $p_2^-(h_{t-1})/p_2^+(h_{t-1})$ with probability $p_2^+(h_{t-1})$, and so forth for those indices $m$ for which $p_m^+ \neq 0$. Consequently, it suffices to show for some $m$,

$$p_m^-(h_{t-1})/p_m^+(h_{t-1}) \leqq 1 - \Delta_0/M \quad \text{and} \quad p_m^+(h_{t-1}) \geqq \Delta_0/M.$$

By hypothesis, $\Delta(h_{t-1}) = \max_m \|p_m^+(h_{t-1}) - p_m^-(h_{t-1})\| \geqq \Delta_0$. Suppose without loss of generality that this maximum occurs at $m = 1$. If $p_1^+(h_{t-1}) - p_1^-(h_{t-1}) \geqq \Delta_0$ then $p_1^+ \geqq \Delta_0$ and $1 - p_1^-(h_{t-1})/p_1^+(h_{t-1}) \geqq \Delta_0/p_1^+(h_{t-1}) \geqq \Delta_0$, so we are done. If, on the other hand, $p_1^-(h_{t-1}) - p_1^+(h_{t-1}) \geqq \Delta_0$, then $\sum_{m>1}(p_m^+(h_{t-1}) - p_m^-(h_{t-1})) \geqq \Delta_0$. Consequently $M \max_{m>1}(p_m^+(h_{t-1}) - p_m^-(h_{t-1}) \geqq \Delta_0$, and, for $m = 2$ say, we have $p_2^+(h_{t-1}) - p_2^-(h_{t-1}) \geqq \Delta_0/M$. Again, we conclude $p_2^+(h_{t-1}) \geqq \Delta_0/M$ and $1 - p_2^-(h_{t-1})/p_2^+(h_{t-1}) \geqq \Delta_0/M$. $\|$

Lemma 4.2 shows that in the periods where the conditional distribution over outcomes induced by the actions of types in $\Omega^- = \Omega\backslash\Omega^+$ differs significantly from that corresponding to $\Omega^+$, the likelihood ratio is likely to jump down by a significant amount. Of course, in periods where $\Delta(h_{t-1})$ is small, the likelihood ratio need not change very much. The key to our result is to show that there is high probability that there are few periods in which both the odds ratio is high and $\Delta(h_{t-1}) > \Delta_0$. To prove this, we introduce a new supermartingale which includes all of the periods where $\Delta(h_{t-1}) > \Delta_0$ from the supermartingale $L(h)$.

We first define a sequence of stopping times relative to a given supermartingale $L = L(h)$ and a distance $\Delta_0$. Set $\tau_0 = 0$. If $\tau_{k-1}(h, \Delta_0) = \infty$, set $\tau_k(h, \Delta_0) = \infty$ as well. If $\tau_{k-1}(h, \Delta_0)$ is finite, set $\tau_k(h, \Delta_0)$ to be the first time $t > \tau_{k-1}(h, \Delta_0)$ such that either

$$\Pr[\|L_t/L_{t-1} - 1\| < \Delta_0/M | h_{t-1}] \geqq \Delta_0/M, \tag{1}$$

or

$$L_t/L_{\tau_{k-1}} - 1 \geqq \Delta_0/2M, \tag{2}$$

or

$$\text{if no such time exists, set } \tau_k(h, \Delta_0) = \infty. \tag{3}$$

Lemma 4.2 shows that this sequence of stopping times picks out at least all the date-history pairs for which $\Delta(h_{t-1}) > \Delta_0$.

The *faster process* $\tilde{L}_k$ relative to $L_t$ and $\Delta_0$ is defined by $\tilde{L}_k = L_{\tau_k}$ for $\tau_k < \infty$, and $\tilde{L}_k = 0$ for $\tau_k = \infty$. Since the $\tau_k$ are stopping times, $\tilde{L}_k$ is a supermartingale, with an associated filtration whose events we denote $\tilde{h}_k$. Since $\tilde{L}_k$ is defined with respect to the stopping times $\tau_k$, the stopping times are measurable with respect to the events $\tilde{h}_{k'}$ for $k' \leqq k$. Consequently, for a fixed $\Delta_0$ and any $k' \leqq k$ the random variable $\tau_k(\tilde{h}_{k'})$ is well defined. Moreover, we will show that $\tilde{L}_k$ is an "active" supermartingale in the following sense:

*Definition.* A positive supermartingale $(\tilde{L}_k, \tilde{h}_k)$ is an active supermartingale with activity $\psi$ if

$$\Pr[\|\tilde{L}_{k+1}/\tilde{L}_k - 1\| > \psi | \tilde{h}_k] > \psi$$

for almost all histories $\tilde{h}_k$ such that $\tilde{L}_k > 0$.

**Lemma 4.3.** *For any $\Omega^+$ with $\mu(\Omega^+) > 0$, and any $\Delta_0 > 0$, the associated faster process $\tilde{L}_k$ is an active supermartingale under $\sigma^+$ with activity $\Delta_0/2M$.*

*Proof.* Since the $\tau_k$ are stopping times, by Lemma 4.1 $\tilde{L}_k$ is a supermartingale. Next, we claim that if $\tilde{h}$ is such that $\tilde{L}_{k-1} > 0$,

$$\Pr[\|\tilde{L}_k/\tilde{L}_{k-1} - 1\| > \Delta_0/2M | \tilde{h}_{k-1}] > \Delta_0/M.$$

To see this, let $s = \tau_{k-1}(\tilde{h})$, which is a constant with respect to $\tilde{h}_{k-1}$; $\tau_k(\tilde{h}_{k-1})$ is a random variable. We will show that

$$\Pr[\|L_{\tau_k}/L_s - 1\| > \Delta_0/2M | \tilde{h}_s] > \Delta_0/M.$$

One of the three rules in the definition of the $\tau$'s must be used to choose $\tau_k$. We will show that this inequality holds conditional on each rule, and thus that it holds averaging over all of them. Conditional on $\tilde{h}_s$, if rule (2) or (3) is used, then with probability one $\|L_{\tau_k}/L_s - 1\| > \Delta_0/2M$. If rule (1) is used, then since inequality (1) holds, and $s = \tau_{k-1} \leqq \tau_k - 1$,

$$Pr[L_{\tau_k}/L_{(\tau_k-1)} - 1 \leqq -\Delta_0/M | \tilde{h}_s, \quad \{\text{rule 1 used}\}] \geqq \Delta_0/M,$$

and also since rule (2) was not used at the date $\tau_k - 1$ just before $\tau_k$,

$$L_{(\tau_k-1)}/L_s - 1 < \Delta_0/2M.$$

Combining the last two inequalities shows that

$$\Pr[L_{\tau_k}/L_s - 1 < -(\Delta_0/2M + \Delta^2/2M^2) \mid \tilde{h}_s, \{\text{rule } 1\}] \geqq \Delta_0/M.$$

Since $-(\Delta_0/2M + \Delta^2/2M^2) < -\Delta_0/2M$, we conclude that

$$\Pr[\|L_{\tau_k}/L_s - 1\| > \Delta_0/2M \mid \tilde{h}_s, \{\text{rule } 1\}) \geqq \Delta_0/M > \Delta_0/2M. \quad \|$$

The remainder of Theorem 4.1 follows from the fact that active supermartingales converge to zero at a uniform rate that depends only on their initial value and their degree of activity.

**Theorem A.1.** *Let $l_0 > 0$, $\varepsilon > 0$, and $\psi \in (0,1)$ be given. For each $\underline{L}$, $0 < \underline{L} < l_0$, there is a time $K < \infty$ such that*

$$\Pr[\sup_{k > K} \tilde{L}_k \leqq \underline{L}] \geqq 1 - \varepsilon$$

*for every active supermartingale $\tilde{L}$ with $\tilde{L}_0 = l_0$ and activity $\psi$.*

This theorem is proved in the Appendix using results about upcrossing numbers. The key aspect of the Theorem is that bound $K$ depends only on $l_0$ and $\psi$, and is independent of the particular supermartingale chosen.

We can now conclude:

**Theorem 4.1.** *For every $\varepsilon > 0$, $\Delta_0 > 0$ and $\Omega^+$ with $\mu(\Omega^+) > 0$ there is a $K$ depending only on these three numbers such that for any $\sigma_1$ and $\sigma_2$, under the probability distribution generated by $\sigma_1(\Omega^+)$, there is probability less than $\varepsilon$ that there are more then $K$ periods with*

$$\|p^+(h_{t-1}) - p(h_{t-1})\| > \Delta_0.$$

*Proof.* Set $\underline{L} = \Delta_0/(1-\Delta_0)$ and $\tilde{L}_0 = (1-\mu(\Omega^+))/\mu(\Omega^+)$. Since by Lemma 4.2 the faster process omits only observations when $\Delta(h_{t-1}) \leqq \Delta_0$, we conclude form Lemma 4.3 and Theorem A.1 that there exists $K$, depending only on $\tilde{L}_0$ and $\Delta_0$, so that with probability $1-\varepsilon$ under $\sigma^+$ in all but $K$ periods either $\Delta(h_{t-1}) \leqq \Delta_0$ or $L_{t-1} \leqq L$. By Lemma 4.1, we conclude $L_{t-1} = [1-\mu(\Omega^+ \mid h_{t-1})]/\mu(\Omega^+(h_{t-1}) \leqq \Delta_0/(1-\Delta_0)$, implying either $\Delta(h_{t-1}) \leqq \Delta_0$ or $\mu(\Omega^+ \mid h_{t-1})) \geqq 1-\Delta_0$. Since $\|p^+(h_{t-1}) - p(h_{t-1})\| \leqq \Delta(h_{t-1})$, the former implies $\|p^+(h_{t-1}) - p(h_{t-1})\| \leqq \Delta_0$, while the latter implies

$$\begin{aligned}\|p^+(h_{t-1}) - p(h_{t-1})\| &= \|p^+(h_{t-1}) - \{\mu(\Omega^+ \mid h_{t-1})p^+(h_{t-1}) + [1-\mu(\Omega^+ \mid h_{t-1})]p^-(h_{t-1})\}\| \\ &\leqq [1-\mu(\Omega^+ \mid h_{t-1})]\|p^+(h_{t-1}) - p^-(h_{t-1})\| \\ &\leqq 1-\mu(\Omega^+ \mid h_{t-1}) \leqq \Delta_0. \quad \|\end{aligned}$$

## APPENDIX: ACTIVE SUPERMARTINGALES

Our goal is to prove

**Theorem A.1.** *Let $l_0 > 0$, $\varepsilon > 0$, and $\psi \in (0,1)$ be given. For each $\underline{L}$, $0 < \underline{L} < l_0$, there is a time $K < \infty$ such that*

$$\Pr[\sup_{k > K} \tilde{L}_k \leqq \underline{L}] \geqq 1 - \varepsilon$$

*for every active supermartingale $\tilde{L}$ with $\tilde{L}_0 = l_0$ and activity $\psi$.*

For a given martingale this is a simple consequence of the fact that $\tilde{L}$ converges to zero with probability one. The force of the theorem is to give a uniform bound on the rate of convergence for all supermartingales with a given activity $\psi$ and initial value $l_0$.

15

Throughout the appendix we use $\tilde{L}$ to denote any supermartingale that satisfies the hypotheses of Theorem A.1. To prove the theorem, we will use some fundamental results from the theory of supermartingales, in particular, bounds on the "upcrossing numbers" which we introduce below. These results can be found in Neveu (1975, Chapter II).

**Fact A.2.** *For any positive supermartingale,* $\Pr[\sup_K \tilde{L}_k \geqq c] \leqq \min(1, \tilde{L}_0/c)$.

Next, fix an interval $[a, b]$, $0 < a < b < \infty$, and define $U_k(a, b)$ to be the number of upcrossings of $[a, b]$ up to time $k$; let $U_\infty(a, b)$ be the total number of upcrossings (possibly equal to $\infty$).

**Fact A.3.** *For any positive supermartingale,* $\Pr[U_\infty(a, b) \geqq N] \leqq (a/b)^N \min(\tilde{L}_0/a, 1)$.

This is known as Dubin's inequality. (See, for example, Neveu (1975, p. 27).)

Next we observe that since $\tilde{L}$ has activity $\psi$, it makes a jump of size $\psi$ with probability at least $\psi$ in each period $k$ where $\tilde{L}_k$ is non-zero. Consequently, over a large number of periods either $\tilde{L}$ has jumped to zero or there are likely to be "many" jumps. Specifically, define $J_k$ to be the number of times $k' < k$ that $\|(\tilde{L}_{k'+1}/\tilde{L}_{k'}) - 1\| > \phi$.

**Lemma A.4.** *For all* $\varepsilon$ *and* $J$ *there exists a* $K$ *such that*

$$\Pr[\{J_K \geqq J\} \text{ or } \{\tilde{L}_K = 0\}] \geqq 1 - \varepsilon.$$

*Proof.* Because $\tilde{L}$ has activity $\psi$, in each period $k'$, either $\tilde{L}_{k'} = 0$ or the probability of a jump of size $\psi$ at time $k'$ exceeds $\psi$. Define a sequence of indicator functions $I_k$ by $I_k = 1$ iff $\{\tilde{L}_k = 0$ or $\|\tilde{L}_k/\tilde{L}_{k-1} - 1\| > \psi\}$, and set $S_k = \sum_{k' \leq k} I_{k'}$. Each $I_k$ has expectation at least $\psi$, so for some $K$ sufficiently large, $\text{Prob}[S_K \geqq J] \geqq 1 - \varepsilon$. Now if $S_K \geqq J$, then either $\tilde{L}_k = 0$ for some $k \leqq K$, in which case $\tilde{L}_K = 0$ as well, or there have been at least $J$ jumps by time $K$. ||

We have now established that most paths of $\tilde{L}$

(1) do not exceed $\bar{c}$ for $\bar{c}$ large, (Fact A.2)

(2) make "few" upcrossing of any positive interval $[a, b]$ (Fact A.3), and

(3) either make "lots of jumps" or hit zero (Lemma A.4).

We will use these three conditions to show that for $K$ large, most paths remain below $\underline{L}$ from $K$ on. To do so, we first argue that most paths will pass below $\underline{c}$ by time $K$.

Divide the interval $[\underline{c}, \bar{c}]$ into $I$ equal sub-intervals with endpoints $e_1 = \underline{c}, \ldots, e_{I+1} = \bar{c}$. Then define the events

$E_1$ if $\max_{k \leq K} \tilde{L}_k \geqq \bar{c}$;

$E_2$ if at least one of the interval $[e_i, e_{i+1}]$ is upcrossed $N$ or more times by time $K$;

$E_3$ if $J_K < J$ and $\tilde{L}_K > 0$,

$E_4$ if $\min_{k \leq K} \tilde{L}_k < \underline{c}$.

By judicious choice of $\bar{c}$, $I$, $K$, $N$ and $J$, we will insure that $E_4^c \subset E_1 \cup E_2 \cup E_3$ and that $\Pr(E_1), \Pr(E_2), \Pr(E_3) \leqq \varepsilon/4$. This will yield our preliminary conclusion that

$$\Pr[\min_{k \leq K} \tilde{L}_k < \underline{c}] = \Pr(E_4) \geqq 1 - (3\varepsilon/4).$$

If we then choose

$$\underline{c} = (\varepsilon/4)\underline{L},$$

Fact A.2 implies that

$$\Pr[\max_{k > K} \tilde{L}_k > \underline{L} \mid \min_{k \leq K} \tilde{L}_k < \underline{c}] \leqq \underline{c}/\underline{L} \leqq \varepsilon/4$$

and we get the desired conclusion that

$$\begin{aligned}\Pr[\max_{k > K} \tilde{L}_k > \underline{L}] &= \Pr[\max_{k > K} \tilde{L}_k > \underline{L} \mid \min_{k \leq K} \tilde{L}_K < \underline{c}] \cdot \Pr[\min_{k \leq K} \tilde{L}_k < \underline{c}] \\ &\quad + \Pr[\max_{k > K} \tilde{L}_k > \underline{L} \mid \min_{k \leq K} \tilde{L}_K \geqq \underline{c}] \cdot \Pr[\min_{k \leq K} \tilde{L}_k \geqq \underline{c}] \\ &\leqq (\varepsilon/4) \cdot 1 + 1 \cdot -(3\varepsilon/4) = \varepsilon\end{aligned}$$

Turning first to $E_1$, we can again use Fact A.2 to choose

$$\bar{c} = (4/\varepsilon) l_0$$

and insure that $\Pr(E_1) = \Pr(\text{Max}_{k \leq K} \hat{L}_k \geqq \bar{c}) \leqq \varepsilon/4$. Note for future reference that this is true, regardless of how we pick $K$.

In the range above $\underline{c}$, when $\|\hat{L}_k/\hat{L}_{k-1} - 1\| > \psi$, $\|\hat{L}_k - \hat{L}_{k-1}\| \geqq \psi\underline{c}$. Thus, if we choose

$$I \geqq 2\bar{c}/\underline{c}\psi + 1$$

then the width of each sub-interval is less than $\psi\underline{c}/2$. This means that each jump of relative size $\psi$ in a path that remains between $\underline{c}$ and $\bar{c}$ must cross one of the sub-intervals $[e_i, e_{i+1}]$. Moreover, if such a path has $J$ or more jumps across sub-intervals, it must cross at least one sub-interval $(J-1)/2I - 1$ times. Consequently if we choose

$$N \leqq (J - I)/2I - 1 \qquad (*)$$

then $E_4^c \subset E_1 \cup E_2 \cup E_3$ as required. In other words, a path that does not go above $\bar{c}$, that does not upcross any subinterval in $[\underline{c}, \bar{c}]N$ or more times, and jumps $K$ or more times, must fall below $\underline{c}$. By Fact A.3, we know that for any given sub-interval, the probability of $N$ or more upcrossings is not more than

$$(1+\psi)^{-N} l_0/\underline{c}.$$

Consequently, the probability that some subinterval is upcrossed $N$ or more times is no more than

$$I(1+\psi)^{-N} l_0/\underline{c}.$$

To make $\Pr(E_2) \leqq \varepsilon/4$ we should choose

$$N \geqq \frac{4Il_0/\underline{c}\varepsilon}{\log(1+\psi)}$$

This determines $J$ by $(*)$ above

$$J = 2I(N+1) + I.$$

Finally, choose $K$ by Fact A.4 to make $\Pr(E_3) \leqq \varepsilon/4$. ||

*Acknowledgement*: This paper benefitted from conversations with Roland Benabou, David Kreps, John McCall, John Riley, Jean Tirole, and especially Shmuel Zamir, who suggested that our results would extend to games with moral hazard. We also thank John Moore and two anonymous referees for helpful comments. The National Science Foundation Grants SES 89-07999, SES 87-08616, SES 86-09697, SES 85-09484, the Sloan Foundation, and the UCLA Academic Senate provided financial support.

REFERENCES

AGHION, P. and CAILLAUD, B. (1988), "On the Role of Intermediaries in Organizations" (mimeo).

AUMANN, R. and SORIN, S. (1989), "Cooperation and Bounded Recall", *Games and Economic Behavior*, **1**, 5-39.

BENABOU, R. and LAROQUE, G. (1988), "Using Privileged Information to Manipulate Markets" (INSEE Working Paper 137/930).

CUKIERMAN, A. and MELTZER, A. (1986), "A Theory of Ambiguity, Credibility and Inflation Under Discretion, and Asymmetric Information", *Econometrica*, **54**, 1099-128.

FUDENBERG, D. and KREPS, D. M. (1987), "Reputation and Simultaneous Opponents", *Review of Economic Studies*, **54**, 541-68.

FUDENBERG, D., KREPS, D. M. and LEVINE, D. K. (1988), "On the Robustness of Equilibrium Refinements", *Journal of Economic Theory*, **44**, 354-80.

FUDENBERG, D., KREPS, D. M. and MASKIN, E. (1990), "Repeated Games with Long-Run and Short-Run Players", *Review of Economic Studies*, **57**, 555-73.

FUDENBERG, D. and LEVINE, D. K. (1983), "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games", *Journal of Economic Theory*, **31**, 251-68.

FUDENBERG, D. and LEVINE, D. K. (1989), "Reputation and Equilibrium Selection in Games with a Single Patient Player", *Econometrica*, **57**, 759-78.

FUDENBERG, D. and MASKIN, E. (1986), "The Folk Theorem in Repeated Games With Discounting or With Incomplete Information", *Econometrica*, **54**, 533-54.

HARSANYI, J. (1967), "Games With Incomplete Information Played by Bayesian Players", *Management Science*, **14**, 159-82 and 320-34.

KREPS, D. M., MILGROM, P., ROBERTS, J. and WILSON, R. (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma", *Journal of Economic Theory*, **27**, 245-52 and 486-502.

KREPS, D. M. and WILSON, R. (1982), "Reputation and Imperfect Information", *Journal of Economic Theory*, **27**, 253–79.

LOÉVE, M. (1978) *Probability Theory II* (New York: Springer-Verlag).

MILGROM, P. and ROBERTS, J. (1982), "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory*, **27**, 280–34.

MILGROM, P. and WEBER, R. (1985), "Distributional Strategies for Games With Incomplete Information", *Mathematics of Operations Research*, **10**, 619–32.

NEVEU, J. (1975) *Discrete-Parameter Martingales* (Amsterdam: North-Holland).

SELTEN, R. (1977), "The Chain-Store Paradox", *Theory and Decision*, **9**, 127–59.