Economica, **56**, 15-24

Reputation and Patience in the 'War of Attrition'

By Lewis Kornhauser,[1] Ariel Rubinstein[2] and Charles Wilson[3]

[1,3] *New York University and* [2] *the Hebrew University*

Final version received 10 February 1988. Accepted 23 February 1988.

The paper presents an approach to selecting among the many subgame-perfect equilibria that exist in a standard concession game with complete information. We extend the description of a game to include a specific 'irrational' (mixed) strategy for each player. Depending on the irrational strategies chosen, we demonstrate that this approach may select a unique equilibrium in which the weaker player concedes immediately. A player is weaker either if he is more impatient or if his irrational strategy is to wait in any period with the higher probability.

I. The War of Attrition

The following game is a variant of the 'War of Attrition' which is now a standard paradigm in economic theory. Some examples of its applications include the study of price wars (e.g. Fudenberg and Tirole, 1986, Ghemawat and Nalebuff, 1985; Kreps and Wilson, 1982), bargaining (Ordover and Rubinstein, 1986; Osborne, 1985), the supply of public goods (Bliss and Nalebuff, 1984), and oil exploration (Wilson, 1982). (See Hendricks and Wilson, 1985, for a survey of the literature.) Two players, 1 and 2, are involved in a dispute. Time is discrete and the players alternately have the option to concede. If player 1 concedes in period t, the outcome is (A, t). If player 2 concedes in period t, the outcome is (B, t). If neither ever concedes, the outcome is (C, ∞).

The game form is illustrated in Figure 1. To enforce the alternation of moves, we restrict player 1 to move only in the even periods and player 2 to move only in the odd periods. A strategy for player 1 then is a sequence $\alpha_1 = \{\alpha_1(t)\}_{t=0,2,4,\ldots}$, where $\{1-\alpha_1(t)\} \in [0, 1]$ is the probability that player 1 concedes in period t conditional on neither player conceding before period t. Similarly, a strategy for player 2 is a sequence $\alpha_2 = \{\alpha_2(t)\}_{t=1,3,5,\ldots}$.

We suppose that the preferences of the players can be represented by Van Neumann-Morgenstern utilities, v_i satisfying $v_1(C, \infty) = v_2(C, \infty) = 0$ and, for $t < \infty$, $v_i(x, t) = u_i(x)\delta_i^t$, where $u_1(A) = u_2(B) = L$ (the low payoff) is the return

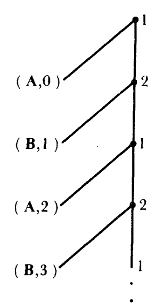

Figure 1. The Extensive Game

to conceding and $u_1(B) = u_2(A) = H$ (the high payoff) is a player's return if the other player concedes. We assume that $0 < \delta_i < 1$ and $H\delta_i > L > 0$ for $i = 1, 2$. Thus, if a player is sure his opponent will concede in the next period, it is optimal for him not to concede, but if he is to be the first to concede, he prefers to do it sooner rather than later.

There are two asymmetries in the model. One is due to the order of the moves in the game. For our purposes, this asymmetry is not important since our results below do not depend on the order in which the players move.[1] The second potential asymmetry is in the time preferences of the players (δ_1 may be different than δ_2.) It is on this asymmetry that we will focus our attention.

Regardless of the relative size of the discount factors, there is an infinity of subgame-perfect equilibrium outcomes. One of these outcomes is for player 1 to concede immediately. Another is for player 1 to wait and for player 2 to concede immediately. Our own intuition, however, suggests that the weaker player, the one with the lower discount factor, should concede immediately. One of the aims of the paper is to develop a criterion for selecting this particular equilibrium outcome.

I. What Is Missing in the Model

We take the position that a game of complete information can generally be thought of as an approximation to a multi-person decision problem in which each player is reasonably certain about the objectives of the other players but does entertain the possibility that one or more of the other players will act irrationally. Consequently, our criterion for selecting an equilibrium will require that it be close to an equilibrium of the corresponding perturbed game obtained by introducing some irrational player with arbitrary small probability. There are, of course, many ways to model an 'irrational' player, and, as Fudenberg and Maskin (1986) have shown in the context of a repeated game, the choice of perturbations may dramatically affect the equilibrium outcome. In this paper we are using perturbations to capture biases in the uncertainty players have about the 'toughness' of their opponents. For the purposes of this paper, therefore, we will identify an irrational player with a particular mixed strategy which we believe captures these biases in beliefs.

Formally, let Γ be an n player game in extensive form and $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ be any (behaviour) strategy combination for Γ. For any $\varepsilon \in (0, 1)$, let $(\Gamma, \sigma^*, \varepsilon)$ be a game with the same extensive form and payoffs as Γ with the property that any strategy combination σ for $(\Gamma, \sigma^*, \varepsilon)$ is equivalent to the strategy combination $\varepsilon\sigma + (1-\varepsilon)\sigma^*$ for Γ.[2] Then, σ is a (sequential) *σ^*-equilibrium* for Γ if it is the limit of a sequence of strategy combinations $\{\sigma^k\}$, each of which is a (sequential) equilibrium for $(\Gamma, \sigma^*, \varepsilon^k)$ for some sequence $\{\varepsilon^k\}$ such that $\varepsilon^k \to 0$.[3] We interpret a σ^*-equilibrium to be the outcome of a game Γ when all of the participants believe that there is some (small) chance that the strategy of player i will deviate from his rational play and adopt strategy σ_i^*.

One may view this approach as the combination of two ideas that are already well established in the literature. In a series of papers, Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982), and Milgrom and Roberts (1982) have included the players' doubts explicitly in the model. In determining the equilibrium outcomes in the Chain Store Paradox and the Prisoner's Dilemma, they suppose that the players assign a small probability that their opponents use a certain strategy specified exogenously by the modeller (the 'tough' Chain Store Strategy and the 'tit-for-tat', accordingly). Under such an assumption, they are able to obtain sequential equilibria which are not equilibria of the original game.

Our objective differs from theirs in two ways. First, we want to determine how the equilibrium changes as we vary the fixed strategy σ^* and/or the payoffs of the rational players. Second, we are interested in selecting an equilibrium from the set of equilibria of the original game of complete information rather than trying to justify a new equilibrium outcome. Thus, although we modify the game by adding an irrational player, we are only interested in the equilibria as the probability of the irrational player goes to zero. This leads us to the second idea to which our idea is related.

Selten (1975), Myerson (1978), Kohlberg and Mertens (1986), and others have suggested equilibrium concepts based on the limit of the equilibria of sequences of perturbed games. The most widely used of these ideas is Selten's concept of a 'trembling hand' perfect equilibrium. An equilibrium is trembling-hand-perfect if it is the limit of the equilibria of some sequence of games in which the behaviour strategy at each information set is perturbed with increasingly small probability. Thus, not only is the specific perturbation unspecified in advance, but any errors across information sets are uncorrelated. Evidently, mistakes are to be interpreted as errors of execution rather than errors of rationality.

The primary motivation behind Selten's approach was to extend the intuition of subgame perfection to games with incomplete information. The motivation behind our approach is to test an equilibrium against a pre-specified possibility of irrational behaviour. Thus, the ideas behind the σ^*-equilibrium differs from the trembling-hand-perfect equilibrium in two ways. First, as in the work on the Chain Store Paradox cited above, our perturbations are in mixed (or behaviour) strategies rather than local strategies (or mixed strategies in the agent-normal form). This leads to the possibility that mistakes are correlated across information sets.[4] Second, it implicitly incorporates a precise form of irrational behaviour (and hence specific perturbations of the strategies) in the concept of the equilibrium. As we emphasized above, our main objective is to study how changes in the form of irrational behaviour affect the outcome of the game.

Finally, we should note that the techniques used to establish our results are also well established in the literature. A number of authors (e.g. Kreps and Wilson, 1982; Fudenberg and Tirole, 1986; and Wilson, 1982) have demonstrated that introducing certain kinds of irrational players in the War of Attrition will lead to unique (or locally unique) equilibria. The form of the irrationality, however, is generally quite simple. The irrational player never concedes. That is, he represents the most extreme version of a 'tough' player. Because we wish to parameterize the 'toughness' of a player and examine its implications for the equilibrium of the game, however, we are forced to require even the irrational players to mix in each period. As a consequence, the *details* of the analysis become considerably more complicated.

We turn now to the concession game described in Section I.

III. THE MAIN RESULT

The theme of our results is illustrated in the following example. Suppose that each player is irrational with probability $\varepsilon > 0$ where for player 2 'irrationality' means never to concede and for player 1 it means to concede always. It is easy to check that, regardless of the values of δ_1 and δ_2, the only equilibrium outcome is for player 1 to concede immediately. This observation fits nicely our intuition that the asymmetry in the content of 'out-of-rationality' behaviour is critical in determining the outcome of the game. Player 2 can build up a reputation of playing tough, whereas player 1 does not have the tools to do that. In this section, we parameterize the ability of players to build up their reputations and investigate its implications for the equilibrium of the game.[5]

In general, there are irrational opponents against whom it is optimal to concede immediately but to wait if the game reaches some later stage without a concession. Consequently, if the influence of an irrational player is to be independent of time, we must impose some stationarity in the strategies of the irrational opponents. We will therefore restrict attention to irrational players whose strategies are of the form $(\gamma, \gamma, \gamma, \ldots)$. That is, the irrational player plans to concede with the same probability $(1-\gamma)$, conditional upon reaching any period in which he is permitted to move. In what follows, let i refer to an arbitrary player and j to the other player, and assume that any integer t is odd or even as the definitions require.

To state our main result, we define p_j to be the solution of

$$\delta_i\{(1-p_j)H + p_j\delta_i L\} = L.$$

Suppose that, conditional on reaching period t, player j concedes with probability $(1-p_j)$. Then player i is indifferent between conceding in period $t-1$ and waiting until period $t+1$ to concede. If, upon reaching period t, player j plans to concede with a probability smaller than $(1-p_j)$, then player i prefers to concede in period $t-1$ rather in period $t+1$. If, upon reaching any period t, player j plans to concede with a probability greater than $(1-p_j)$, then player i prefers to wait until period $t+1$ rather than concede in period $t-1$. Since we suppose that $\delta_i H > L$, it follows that $0 < p_j < 1$. Furthermore, $\delta_i > \delta_j$ implies $p_j > p_i$.

Let $\Gamma(\delta_1, \delta_2)$ be the concession game defined by the discount factors δ_1 and δ_2, and let $(\gamma_1, \gamma_2) \in [0, 1]^2$ represent a pair of stationary strategies. Then, given the definition of p_i, we may state our main result as follows.

Theorem 1. Suppose $\gamma_2 > p_2$ and $\gamma_2/p_2 > \gamma_1/p_1$. Then the unique sequential (γ_1, γ_2)-equilibrium outcome for $\Gamma(\delta_1, \delta_2)$ is for player 1 to move immediately with probability 1 (i.e., $\alpha_1(0) = 0$).

Recall that $(1-\gamma_i)$ is the probability that, upon reaching any period, the irrational player i concedes, and $(1-p_i)$ is the probability of moving in any period that induces indifference for player j between immediate concession and waiting to concede at his next turn. If γ_2 is greater than p_2, then, faced with his irrational opponent, player 1 would concede immediately. If in addition the ratio γ_2/p_2 is greater than γ_1/p_1, then Theorem 1 implies that player 1 concedes immediately in any sequential (γ_1/γ_2)-equilibrium. In particular, if the players have identical time preferences $(\delta_1 = \delta_2)$, the player with the better facility for building a reputation for toughness (the highest γ_i) will win, while if the players have the same facilities for establishing a reputation ($\gamma_1 = \gamma_2$), then the more impatient player concedes immediately.

Theorem 1 is a statement about the equilibrium outcomes. For almost all parameter values, the sequential (γ_1, γ_2)-equilibrium is itself unique.

Theorem 2. (a) Suppose $\gamma_2/p_2 > \gamma_1/p_1 > 1$. Then $\alpha_1 = (0, p_1, p_1, p_1, \ldots)$ and $\alpha_2 = (p_2, p_2, p_2, \ldots)$ is the unique sequential (γ_1, γ_2)-equilibrium for $\Gamma(\delta_1, \delta_2)$.

(b) Suppose $\gamma_2/p_2 > 1 > \gamma_1/p_1$. Then $\alpha_1 = (0, 0, 0, \ldots)$ and $\alpha_2 = (1, 1, 1, \ldots)$ is the unique sequential (γ_1, γ_2)-equilibrium for $\Gamma(\delta_1, \delta_2)$.

Given the conditions of Theorem 1, Theorem 2 reveals a kind of second-order benefit to player 1 if his irrational counterpart (who plays γ_1) is sufficiently tough. When $\gamma_1/p_1 < 1$, player 2 always waits and player 1 always concedes, regardless of the history of the game. However, when $\gamma_1/p_1 > 1$, each player i concedes with probability $(1 - p_i)$ upon reaching any later period. Thus, if player 1 makes a 'mistake' in the first period and waits, there is positive probability that player 2 will eventually concede.

Theorem 1 depends upon the satisfaction of two conditions. First, at least one of the players must have the ability to build a reputation for toughness. Second, one of the players must have an advantage over his opponent in building his reputation. If either of these conditions is violated, we obtain a different set of sequential (γ_1, γ_2)-equilibrium outcomes. If $\gamma_1/p_1 = \gamma_2/p_2 > 1$, then both players have an equal facility for building a reputation for toughness. In this case, player 1 concedes immediately with a probability between $1 - p_1$ and $1 - \gamma_1$. Thereafter, the probability with which player 1 concedes depends only on the impatience of the other player.[6] On the other hand, if neither player is sufficiently tough to induce a rational opponent to concede, then it is a sequential (γ_1, γ_2)-equilibrium outcome for either player to concede immediately.[7]

If we reverse the order of γ_1/p_1 and γ_2/p_2, the statement of the theorems must be modified, but the results are essentially the same.

IV. Proofs

In the section, we provide a complete proof of Theorem 2(a). The proof of Theorem 2(b) (and Theorem 1, for the case where $\gamma_1 = p_1$) is provided in an earlier version of the paper.

Suppose player 1 is playing strategy α_1. We will use the following notation. For any odd period t, let $\mu_1(t)$ be the probability player 2 assigns to the possibility that player 1 is an irrational player, conditional on the game reaching period t. Let $1 - \beta_1(t+1)$ be the probability player 2 assigns to the possibility that player 1 plans to concede in period $t+1$ conditional on reaching period $t+1$. Then, letting $\mu_1(-1) = \varepsilon$, we may define, for any odd period $t > 0$,

$$\mu_1(t+2) = \gamma_1 \mu_1(t)/\beta_1(t+1)$$

and

$$\beta_1(t+1) = \{1 - \mu_1(t)\}\alpha_1(t+1) + \mu_1(t)\gamma_1.$$

We begin by characterizing the equilibria of the perturbed game $\{\Gamma(\delta_1, \delta_2), (\gamma_1, \gamma_2), \varepsilon\}$. The argument is roughly as follows. We first establish that, as long as the rational player i has not yet moved with probability 1, he must adjust his strategy so that (after the initial period) his opponent is indifferent between conceding immediately and waiting another period. Since $\gamma_i > p_i$, this implies that, in any period t, the rational player j concedes with a higher probability than his irrational counterpart. Consequently, we eventually reach a period $\hat{t}_i$ by which he has conceded with probability 1. At this point, faced with a relatively 'tough' irrational opponent, it is optimal for the rational player j to concede immediately if he has not already done so. This argument leads to the conclusions of Lemma 2 below.

In Lemmas 1-4, assume $\gamma_2/p_2 > \gamma_1/p_1 > 1$ and take $\beta_i(t)$, $\mu_i(t)$, etc., to be equilibrium values for player i in the game $\{\Gamma(\delta_1, \delta_2), (\gamma_1, \gamma_2), \varepsilon\}$.

Lemma 1. Suppose $t > 0$. Then, for $i = 1, 2$,

(a) $\beta_i(t+1) \geq p_i$, and

(b) $\beta_i(t) > p_i$ implies $\alpha_i(t) = 0$ (and hence $\mu_i(t+1) = 1$).

Proof. We will establish part (b) first. Suppose $\beta_i(t) > p_i$ for some $t > 0$. Then, if it is optimal for the rational player j to wait in period $t-1$, it cannot be optimal for him to move in period $t+1$. Therefore, the rational player j either moves with certainty upon reaching period $t-1$ or waits with certainty upon reaching period $t+1$. In the first instance, player j is irrational with probability 1 in period $t+1$ so that $\beta_j(t+1) = \gamma_j$. In the second instance, $\beta_j(t+1) = \{1 - \mu_j(t)\} + \mu_j(t)\gamma_j \geq \gamma_j$. In either instance, $\beta_j(t+1) \geq \gamma_j > p_j$.

Proceeding by induction, it follows that $\beta_j(t+k) > p_j$ for all odd $k > 0$. Consequently, for every non-negative even k, upon reaching period $t+k$, the rational player j strictly prefers conceding immediately to waiting and conceding in period $t+k+2$. But since the game is continuous at infinity,[8] it follows that, upon reaching period t, the rational player j strictly prefers conceding in period t in waiting until any later period (including ∞). Therefore, $\alpha_i(t) = 0$ and hence $\mu_i(t+1) = 1$.

To establish part (a), suppose $\beta_i(t+1) < p_i$ for some $t > 0$. Then, upon reaching period t, the rational player j will choose to wait; i.e., $\alpha_j(t) = 1$. Therefore $\beta_j(t) = \{1 - \mu_j(t-1)\} + \mu_j(t-1)\gamma_j \geq \gamma_j > p_j$, contradicting part (b). Q.E.D.

For $i = 1, 2$, define $\hat{t}_i \equiv \sup\{t : \mu_i(t) < 1\}$ to be the last period in which player j moves for which there is still a positive probability that player i is rational (or ∞).[9] Then we may establish

Lemma 2. (a) $\beta_i(t) = p_i$ for $2 \leq t < \hat{t}_i$, $i = 1, 2$; (b) $\hat{t}_i < \infty$ for $i = 1, 2$; and (c) $|\hat{t}_2 - \hat{t}_1| = 1$.

Proof. Part (a) follows immediately from Lemma 1 and the definition of $\hat{t}_i$. To establish (b), suppose that $\hat{t} = \infty$. Then it follows from part (a) that, for all $t > 1$, $\beta_i(t) = p_i$ and hence that $\mu_i(t) = (\gamma_i/p_i)\mu_i(t-2)$. But then, $\gamma_i/p_i > 1$ implies that $\mu_i(t) > 1$ for t sufficiently large: a contradiction.

To establish (c), suppose that $\mu_i(t) = 1$. Then, for all odd $k > 0$, $\beta_i(k+t) = \gamma_i > p_i$. The lemma then follows from part (a) of Lemma 1. Q.E.D.

The next step is to note that $\gamma_2/p_2 > \gamma_1/p_1$ implies that the rate at which rational player 1 must concede in order to make his opponent indifferent between conceding and waiting is smaller than for player 2. Consequently, to ensure that the rational players first concede with probability 1 in adjacent periods, player 1 must concede with a relatively high probability at the outset of the game. Furthermore, the smaller is the initial probability that the players are irrational, the longer the rational players must wait before conceding with probability 1, and consequently, the larger is the probability that player 1 must concede at the outset in order to ensure that both players first concede with certainty in adjacent periods. Theorem 2(a) is then proved by establishing that, in the limit, this relationship requires player 1 to concede immediately with probability 1.

Lemma 3. (a) $\hat{t}_1 = \hat{t}_2 - 1$ implies that

$$\mu_2(2)(\gamma_2/p_2)^{(\hat{t}_2-2)/2} < 1 = \mu_1(1)(\gamma_1/p_1)^{\hat{t}_2/2} \leq \mu_2(2)(\gamma_2/p_2)^{\hat{t}_2/2};$$

(b) $\hat{t}_1 = \hat{t}_2 + 1$ implies that

$$\mu_1(1)(\gamma_1/p_1)^{\hat{t}_2/2} < 1 = \mu_2(2)(\gamma_2/p_2)^{\hat{t}_2/2} \leq \mu_1(1)(\gamma_1/p_1)^{(\hat{t}_2+2)/2}.$$

Proof. We establish first that $\hat{t}_j = \hat{t}_i + 1$ implies $\beta_i(\hat{t}_i + 1) = p_i$. Suppose $\hat{t}_j = \hat{t}_i + 1$ but $\beta_i(\hat{t}_i + 1) \neq p_i$. Then Lemma 1(a) implies that $\beta_i(\hat{t}_i + 1) > p_i$. Consequently, upon reaching period $\hat{t}_i$, the rational player j strictly prefers conceding immediately to waiting and moving in period $\hat{t}_i + 2$. Therefore, either $\alpha_j(\hat{t}_i) = 0$ or $\alpha_j(\hat{t}_i + 2) = 1$. In the first case, $\mu_j(\hat{t}_i) = 1$, violating the definition of $\hat{t}_i$, and in the second case, $\mu_j(\hat{t}_j + 2) < \mu_j(\hat{t}_j) < 1$, violating the definition of $\hat{t}_j$.

Now suppose $\hat{t}_1 = \hat{t}_2 - 1$. Then the previous paragraph implies that $\beta_i(\hat{t}_i + 1) = p_i$ and hence that

$$\mu_2(2)(\gamma_2/p_2)^{(\hat{t}_2-2)/2} = \mu_2(\hat{t}_2) < 1 = \mu_1(\hat{t}_2 + 2) = \mu_1(1)(\gamma_1/p_1)^{\hat{t}_2/2}.$$

By assumption, $\hat{t}_2 > 0$, and therefore Lemma 1 implies that $\beta_2(\hat{t}_2 + 1) \geq p_2$. Consequently,

$$\begin{aligned} 1 &= \mu_2(\hat{t}_2 + 2) \\ &= \mu_2(2)(\gamma_2/p_2)^{(\hat{t}_2-2)/2}\{\gamma_2/\beta_2(\hat{t}_2 + 1)\} \\ &\leq \mu_2(2)(\gamma_2/p_2)^{\hat{t}_2/2}. \end{aligned}$$

These two relations establish part (a). A similar argument establishes part (b). Q.E.D.

Using Lemma 3, we may establish some properties of the sequential equilibria of $\{\Gamma(\delta_1, \delta_2), (\gamma_1, \gamma_2), \varepsilon\}$ as ε becomes small.

Lemma 4. For any $t > 0$ and any $\varepsilon_1 > 0$, there is a $\psi > 0$ such that, for all $\varepsilon < \psi$: (a) $\beta_2(1) = p_2$, (b) $\hat{t}_i > t$, and (c) $|\alpha_i(t) - p_i| < \varepsilon_1$.

Proof. We establish first that $\beta_2(1) = p_2$ for $\varepsilon > 0$ sufficiently small. Suppose $\beta_2(1) < p_2$. Then $\mu_2(2) > \mu_2(0)\gamma_2/p_2 = \varepsilon\gamma_2/p_2$. Furthermore, player 1 never concedes in period 0, and, therefore, is more likely to be rational in period 1 than at the outset of the game. That is, $\mu_1(1) < \mu_1(-1) = \varepsilon$. Then, using both parts (a) and (b) of Lemma 3, we may show that

$$\varepsilon(\gamma_2/p_2)^{\hat{t}_2/2} < \mu_2(2)(\gamma_2/p_2)^{(\hat{t}_2-2)/2} \leq \mu_1(1)(\gamma_1/p_1)^{\hat{t}_2/2} < \varepsilon(\gamma_1/p_1)^{\hat{t}_2/2},$$

contradicting the assumption that $\gamma_1/p_1 \leq \gamma_2/p_2$. We conclude that $\beta_2(1) \geq p_2$. This implies in turn that $\mu_2(2) \leq \varepsilon\gamma_2/p_2 < 1$ for ε sufficiently small. Part (a) then follows from Lemma 1(b).

Since $\beta_2(1) = p_2$, Lemma 3 implies that

$$\varepsilon(\gamma_2/p_2)^{(\hat{t}_2+2)/2} = \mu_2(2)(\gamma_2/p_2)^{\hat{t}_2/2} \geq 1.$$

Fix $t > 1$. Then there is an $\psi > 0$ such that $\varepsilon < \psi$ implies $t < \hat{t}_2$. Part (b) then follows from Lemma 2(c). Furthermore, it follows from Lemma 2(a) that $\beta_1(k) = p_1$ for even $k, 0 < k \leq t$. Lemma 3 then yields

$$\begin{aligned} \mu_1(t) &= \mu_1(1)(\gamma_1/p_1)^{(t-1)/2} \\ &= \mu_1(1)(\gamma_1/p_1)^{(\hat{t}_1-1)/2}(p_1/\gamma_1)^{(\hat{t}_1-t)/2} \\ &< (p_1/\gamma_1)^{(\hat{t}_1-t)/2}. \end{aligned}$$

Since $\hat{t}_1 \to \infty$ as $\varepsilon \to 0$, it follows that $\mu_1(t) \to 0$ as $\varepsilon \to 0$. Then since $\beta_1(t) = p_1$, it follows from the definition of $\beta_1(t)$ that $\alpha_1(t) \to p_1$ as $\varepsilon \to 0$. A similar argument establishes that, for any odd $t > 0$, $\alpha_2(t) \to p_2$ as $\varepsilon \to 0$. Q.E.D.

Proof of Theorem 2(a). For each $\varepsilon > 0$, it is straightforward to construct a sequential equilibrium satisfying the conditions of Lemmata 3 and 4. (Alternatively, see Fudenberg and Levine, 1986.)

Using the definitions of $\mu_1(1)$ and $\beta_1(0)$ and both parts of Lemma 3, we obtain

$$\begin{aligned} \{\varepsilon\gamma_1 + (1-\varepsilon)\alpha_1(0)\}/\gamma_1 = \beta_1(0)/\gamma_1 = \varepsilon/\mu_1(1) &= (p_2/\gamma_2)\mu_2(2)/\mu_1(1) \\ &\leq (\gamma_1 p_2/\gamma_2 p_1)^{(\hat{t}_2+2)/2}. \end{aligned}$$

Then, since $\gamma_2/p_2 > \gamma_1/p_1$ and $\hat{t}_2 \to \infty$ as $\varepsilon \to 0$ (by Lemma 4(b)), it follows that $\alpha_1(0) \to 0$ as $\varepsilon \to 0$. Lemma 4 also implies that $\alpha_i(t) \to p_i$ for $t > 0, i = 1, 2$. By definition, this strategy pair forms a sequential (γ_1, γ_2)-equilibrium. Q.E.D.

The precise proof of part (b) of Theorem 2 is a bit more cumbersome. Roughly, the idea is to show that, even in the perturbed game, the rational player 1 must move with probability 1 at the outset of the game. Suppose not. Then, since the irrational player 1 concedes with a relatively high probability in any period, the rational player 1 must concede with a lower probability in order to make the rational player 2 indifferent to conceding and waiting. This implies that the rational player 1 never concedes with probability 1. However, to make the rational player 1 indifferent between moving and waiting in each period, an argument similar to that given above implies that the rational player 2 must concede with probability 1 in some period. At this point, faced with the relatively tough irrational player 2, it is optimal for the rational player 1 also to concede immediately. This contradiction establishes the result. A complete proof is given in the unpublished version of this paper.

V. Concluding Remarks

In this paper, we have illustrated an approach that, in some cases, allow us to select a specific outcome in games with multiple equilibria. Essentially, we are suggesting that the specification of the equilibrium be extended to reflect the details about the kind of response a player expects to face if his opponent deviates from rational behaviour. For the particular concession game analysed here, we have parameterized how the differences in the tendency of players to play excessively tough (or weak) affects the interaction of rational players. In an earlier version of the paper, we provided a more complete analysis of how the equilibria depend on the characteristics of the irrational opponent. There we demonstrated that, when the expected irrational behaviour of both players is excessively weak, this approach yields no additional restrictions on the equilibrium outcomes. Only when at least one irrational opponent is sufficiently 'tough' does our concept leads to a unique equilibrium outcome.

In closing, we should emphasize again that our motivation is very different from writers trying to develop a criterion for rational responses to unanticipated events. Indeed, most of the equilibria we eliminate imply that no event can be completely unanticipated. Rather, our objective is to investigate how different presumptions about the irrational tendencies of an opponent can affect rational play. Thus, we are not interested here in refining the equilibrium concept. Nor are we interested in defining a single set of criteria for selecting an equilibria, since we believe that the determinants of equilibrium selection lie outside the formal definition of an extensive game. Although our approach may seem less satisfying than using more rigid criteria, we believe it is preferable to make explicit the presumptions we have in certain situations rather than obscure them behind artificial criteria, the motivation of which is somewhat vague.

ACKNOWLEDGEMENTS

Support from the Filomen D'Agostino and Max E. Greenberg Research Fund of the New York University School of Law is gratefully acknowledged by Lewis Kornhauser. Support from the Alfred P. Sloan Foundation, the C.V. Starr Center for Applied Economics and the Institute for Advanced Studies at the Hebrew University in Jerusalem is gratefully acknowledged by Charles Wilson.

NOTES

1. This asymmetry could be eliminated by supposing that the players move simultaneously in each period. Our general results remain unchanged, but the analysis of the equilibrium becomes more complicated.
2. If $\Pi_i(\sigma)$ is the payoff to player i from strategy combination σ for the game Γ, then $\Pi_i^*(\sigma) = \Pi_i\{\varepsilon\sigma + (1-\varepsilon)\sigma^*\}$ is the payoff to player i in the game $(\Gamma, \sigma^*, \varepsilon)$ from strategy combination σ_-.
3. We use the product topology induced by the Euclidean norm on the space of local strategies (see Fudenberg and Levine, 1983). Given any strategy combination σ^*, the existence of a (sequential) σ^*-equilibrium for games with a finite strategy space follows from standard arguments. The results of Fudenberg and Levine can also be used to guarantee the existence of a (sequential) σ^*-equilibrium for an important class of infinite-horizon games, including the one we study here. The equilibrium concept could obviously be used with other refinements of the equilibrium concept besides sequential equilibrium.
4. See Binmore (1987) for a discussion about the relation between correlated trembles and irrational behaviour.
5. For a discussion of this use of the concept of reputation, see Wilson (1985).
6. This result is established in the unpublished version of the paper.
7. It is enough to verify that (i) $\alpha_1 = (0, 0, 0, \dots)$ and $\alpha_2 = (1, 1, 1, \dots)$ and (ii) $\alpha_1 = (1, 1, 1, \dots)$ and $\alpha_2 = (0, 0, 0, \dots)$ are both sequential (γ_1, γ_2)-equilibria when $\gamma_2/p_2, \gamma_1/p_1 \le 1$. In fact, we can show that, under this assumption, the set of sequential (γ_1, γ_2)-equilibria for $\{\Gamma(\delta_1, \delta_2), (\gamma_1, \gamma_2)\}$ is equal to the set of subgame perfect equilibria for $\Gamma(\delta_1, \delta_2)$.
8. See Fudenberg and Levine (1983) for a precise definition.
9. If $\hat{t}_i < \infty$, then $\mu_i(\hat{t}_i + 2) = 1$.

REFERENCES

BINMORE, K. (1987). Modeling rational players: Part I. *Economics and Philosophy*, **3**, 179-214.

BLISS, C. and NALEBUFF, B. (1984). Dragon-slaying and ballroom dancing: the private supply of a public good. *Journal of Public Economics*, **25**, 1-12.

FUDENBERG, D. and LEVINE, D. (1983). Subgame-perfect equilibria of finite- and infinite-horizon games. *Journal of Economic Theory*, **31**, 251-68.

FUDENBERG, D. and MASKIN, E. (1986). The Folk Theorem for repeated games with discounting or with incomplete information. *Econometrica*, **54**, 533-54.

FUDENBERG, D. and TIROLE, J. (1986). A theory of exit in oligopoly. *Econometrica*, **54**, 943-60.

GHEMAWAT, R. and NALEBUFF, B. (1985). Exit. *Rand Journal of Economics*, **16**, 184-94.

HENDRICKS, K. and WILSON, C. (1985). The war of attrition in discrete time. C.V. Starr Center Economic Research Report no. 85-32.

KOHLBERG, E. and MERTENS, J. F. (1986). On the Strategic Stability of Equilibria. *Econometrica*, **54**, 1003-38.

KORNHAUSER, L., RUBINSTEIN, A., and WILSON, C. (1986). 'Reputation and Patience in the "War of Attrition".' C.V. Starr Center Discussion Paper no. 86-06.

KREPS, D., MILGROM, P., ROBERTS, J., and WILSON, R. (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory*, **27**, 245-52.

KREPS, D. and WILSON, R. (1982). Reputation and imperfect competition. *Journal of Economic Theory*, **27**, 253-79.

MYERSON, R. (1978). Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, **4**, 73-80.

MILGROM, P. and ROBERTS, J. (1982). Predation, Reputation, and Entry Deterrence. *Journal of Economic Theory*, **27**, 280-312.

ORDOVER, J. A. and RUBINSTEIN, A. (1986). 'A sequential concession game with asymmetric information'. *Quarterly Journal of Economics*, **51**, 879-88.

OSBORNE, M. J. (1985). The role of risk aversion in a simple bargaining model. In A. E. Roth (ed.), *Game-theoretic Models of Bargaining*. Cambridge University Press.

SELTEN, R. (1975). A re-examination of the perfectness concept for equilibrium points in extensive games., *International Journal of Game Theory*, **4**, 25-55.

WILSON, C. (1982). Games of timing with incomplete information: an application to oil exploration. Unpublished manuscript, New York University.

WILSON, R. (1985). Reputations in games and markets. In A. Roth (ed.), *Game-theoretic Models of Bargaining*. Cambridge University Press.