ECONOMIC RESEARCH REPORTS

REPUTATION AND PATIENCE IN THE "WAR OF ATTRITION"

by

Lewis Kornhauser,
Ariel Rubinstein
and
Charles Wilson

R.R. #86-06 February 1986

C. V. STARR CENTER
FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

Reputation and Patience in the "War of Attrition"

by

Lewis Kornhauser*
New York University

Ariel Rubinstein**
The Hebrew University

and

Charles Wilson***
New York University

February 1986

* Support from the Filomen D'Agostino and Max E. Greenberg Research Fund of the New York University School of Law is gratefully acknowledged.

** Part of the author's research was completed during his military reserve service in Lebanon. He would like to thank the Israeli Defense Forces for their gracious hospitality.

*** 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.

Abstract

The paper presents an approach to selecting among the many subgame perfect equilibria which 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. An "extended equilibrium" is then an equilibrium of the original game that is the limit of an equilibrium of the corresponding game of incomplete information obtained by introducing the irrational strategies with arbitrary small probability. Depending on the irrational strategies chosen, we demonstrate that, for the "war of attrition", this approach may select a unique equilibrium in which the weaker player moves immediately. A player is weaker if either he is more impatient or his irrational strategy is to wait in any period with the higher probability.

1. The War of Attrition

The following game is a variant of the "War of Attrition" which is now a standard paradigm in economic theory. (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 is then a sequence $\alpha_1 = (\alpha_1(t))_{t=0,2,4,\ldots}$, where $\alpha_1(t) \in [0,1]$ is the probability that player 1 will not concede in period t conditional on the game reaching 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 VNM 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 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 move immediately. Another is for player 1 to wait and for player 2 to move 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.

2. 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. There are, of course, many ways to model an "irrational" player. For the purposes of this paper, we will identify an irrational player with a particular mixed strategy. Our criterion for selecting an equilibrium will then require that it be close to an equilibrium of the corresponding perturbed game obtained by introducing some irrational player with arbitrary small probability.

****

[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.

One may view our 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 in 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 approach differs from theirs in that 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 (1985), 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 behavior 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 prespecified possibility of irrational behavior. Thus, our concept of an "extended" equilibrium differs from the trembling hand perfect equilibrium in two ways. First, our perturbations are in mixed 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.[2] Second, we specify as part of the description of the "extended" game the precise form of irrational behavior (and hence the specific perturbations of the strategies) to be considered.

3. The Concept of an Extended Game

For any n player game Γ in extensive form and any vector of behavior strategies $\underline{\sigma}^* = (\sigma_1^*, \ldots, \sigma_n^*)$, we will call $(\Gamma, \underline{\sigma}^*)$ an <u>extended game</u>. We interpret the extended game as a statement of the underlying game plus the beliefs of the players about the form of the possible irrationality of the others. Each player believes that player i has the preferences defined by Γ with probability near 1 but believes that there is some chance that the strategy of player i is σ_i^*.

For any $\epsilon \in (0,1)$, let $(\Gamma, \underline{\sigma}^*, \epsilon)$ be a game with the same extensive form and payoffs as Γ with the property that any strategy combination $\underline{\sigma}$ for $(\Gamma, \underline{\sigma}^*, \epsilon)$ is equivalent to the strategy combination $\epsilon\underline{\sigma}+(1-\epsilon)\underline{\sigma}^*$ for Γ.[3]

****

[2] See Binmore (1985) for a discussion about the relation between correlated trembles and irrational behavior.

[3] If $\Pi_i(\underline{\sigma})$ is the payoff to player i from strategy combination $\underline{\sigma}$ for the

Then, given some metric function d on the space of strategies, we may define an <u>extended</u> (subgame perfect) <u>equilibrium</u> for $(\Gamma,\underline{\sigma}^*)$ as a strategy combination $\underline{\sigma}$ such that, for any $\epsilon > 0$, there is an $\epsilon_1 \in (0,\epsilon)$ and a strategy combination $\underline{\sigma}'$ such that (i) $d(\sigma_i,\sigma_i') < \epsilon$, $i = 1,\ldots,n$, and (ii) $\underline{\sigma}'$ is a (subgame perfect) equilibrium for $(\Gamma,\underline{\sigma}^*,\epsilon_1)$.

For games with a finite strategy space, the existence of an extended equilibrium (using the Euclidean norm) follows from standard arguments.[4]

With these concepts in hand, we turn to the concession game described in Section 1.

4. The Main Result

We will use the concept of an extended subgame perfect equilibrium to study how the specification of beliefs about the form of the possible irrationality of the other player influences the equilibrium outcome of the concession game described in Section 1. Our basic point can be illustrated by the following example. Suppose that each player is irrational with probability $\epsilon > 0$ 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" behavior is critical in determining the outcome of the game. Player 2 can build up a reputation of

****

game Γ, then $\Pi_i^*(\underline{\sigma}) = \Pi_i(\epsilon\underline{\sigma} + (1-\epsilon)\underline{\sigma}^*)$ is the payoff to player i in the game $(\Gamma,\underline{\sigma}^*,\epsilon)$ from strategy combination $\underline{\sigma}$.

[4] An equilibrium $\sigma(\epsilon)$ exists for each $(\Gamma,\underline{\sigma}^*,\epsilon)$. Letting $\epsilon \to 0$, we may extract a convergent subsequence. The limit is an extended equilibrium for $(\Gamma,\underline{\sigma}^*)$. Note that it is also an equilibrium for Γ.

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 $\underline{\gamma} = (\gamma, \gamma, \gamma, \ldots)$. 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. This leads to the class of extended games $(\Gamma(\delta_1, \delta_2), (\gamma_1, \gamma_2))$ where δ_1 and δ_2 define the original game $\Gamma(\delta_1, \delta_2)$ and $(\underline{\gamma}_1, \underline{\gamma}_2)$ are stationary strategies of the irrational players 1 and 2.

In what follows, let i refer to an arbitrary player and j to the other player. Assume any integer t is odd or even as the definitions require.

To state our main result, let p_j be the solution of the equation

$$\delta_i[(1-p_j)H + p_j\delta_i L] = L.$$

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, conditional on reaching period t, player j plans to wait with a probability greater than p_j, then

****

[5] For a discussion of this use of the concept of reputation, see Wilson (1985).

player i prefers to concede in period t–1 rather in period t+1. If, conditional on reaching any period t, player j plans to wait with a probability less than 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 $p_j > p_i$.

Finally, let d be any metric which generates the product topology in the space of strategies. For instance, we may define the distance between two strategies α_1 and α_1' as

$$d(\alpha_1,\alpha_1') = \sum_{t=0}^{\infty} |\alpha_1(2t)-\alpha_1'(2t)| 2^{-t}.$$

The distance between α_2 and α_2' may be defined similarly. Then we may state our main result as follows.

<u>Theorem 1</u>: Consider $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2))$ with the metric d. Suppose $\gamma_2 > p_2$. Then $\gamma_2/p_2 > \gamma_1/p_1$ implies that $\alpha_1(0) = 0$ is the unique extended subgame perfect equilibrium outcome.

Recall that γ_i is the probability that, upon reaching any period, the irrational player i does not concede, and p_i is the probability of waiting 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 extended subgame perfect 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 extended subgame perfect equilibrium is itself unique.

<u>Theorem 2</u>: Consider $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2))$ with the metric d.

(a) If $\gamma_2/p_2 > \gamma_1/p_1 > 1$, then $\alpha_1 = (0,p_1,p_1,p_1,\ldots)$ and $\alpha_2 = (p_2,p_2,p_2,\ldots)$ is the unique extended subgame perfect equilibrium.

(b) If $\gamma_2/p_2 > 1 > \gamma_1/p_1$, then $\alpha_1 = (0,0,0,\ldots)$ and $\alpha_2 = (1,1,1,\ldots)$ is the unique extended subgame perfect equilibrium.

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 $\gamma_1/p_1 > 1$, either 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 are violated, we obtain a different set of extended subgame perfect equilibrium outcomes.

Theorem 3: Consider $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2))$ with the metric d.

(a) If $\gamma_2/p_2 = \gamma_1/p_1 > 1$, then (α_1,α_2) is an extended subgame perfect equilibrium if and only if $\alpha_1 = (\alpha_1(0),p_1,p_1,p_1,\ldots)$ and $\alpha_2 = (p_2,p_2,p_2,\ldots)$, where $\alpha_1(0) \in [p_1,\gamma_1]$.

(b) If $\gamma_2/p_2,\gamma_1/p_1 \leq 1$, then (i) $\alpha_1(0) = 0$ and (ii) $\alpha_1(0) = 1$ and $\alpha_2(1) = 0$ are both extended subgame perfect 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. If neither irrational player is sufficiently tough to induce a rational opponent to concede, then it is an extended subgame perfect equilibrium outcome for either player to concede immediately.[6]

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.

5. Concluding Remarks

In this paper, we have developed some concepts which, in some cases, allow us to select a specific outcome in games with multiple equilibria. Essentially, we are suggesting that the specification of a game be extended to include the details about the kind of response a player expects to face if his

****

[6] In fact, we can show that when $\gamma_i/p_i \leq 1$ for $i = 1,2$, the set of extended subgame perfect equilibria for $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2))$ is equal to the set of perfect equilibria for $\Gamma(\delta_1,\delta_2)$.

opponent deviates from rational behavior. For the particular concession game we have examined above, we have parameterized how the differences in the tendency of players to play excessively tough (or weak) affects the interaction of rational players.

As we have seen, for some parameter values (i.e. when the irrational behavior of both players tends to be excessively weak), this approach yields no additional restrictions on the equilibrium outcomes. For other parameter values, however, our concept leads to a unique equilibrium outcome.[7] Philosophically, this approach is very different from the approach of many other writers, including Kohlberg-Mertens who seek a single criterion which all games must satisfy. 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.

****

[7] The possibility that the choice of irrational perturbations may affect the equilibrium outcome is illustrated most dramatically in a recent paper by Fudenberg and Maskin (1986). In the context of a repeated game, they show that every individually rational payoff can be approximated as an equilibrium payoff of a game with the proper choice of irrational behavior.

Appendix

In this appendix, we establish a series of lemmata which lead to the the proofs of Theorems 1 to 3. To establish our results, we study in Lemmata 1 to 8 the structure of the game $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ for small values of ϵ. The proofs of the theorems then follow from a study of the equilibria of $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ as ϵ is made arbitrarily small.

To simplify the exposition, we will suppose that $1 > \gamma_i > 0$ for $i = 1,2$.[8]

Suppose player 1 is playing strategy α_1. We will use the following notation. Let $\mu_1(t)$ be the probability that player 2 believes that player 1 is an irrational player conditional on the game reaching period t. Define $\mu_1(-1) = \epsilon$, and, for any odd period t,

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

where

$$\beta_1(t+1) = [1-\mu_1(t)]\alpha_1(t+1) + \mu_1(t)\gamma_1$$

is the probability that player 2 believes that player 1 will <u>not</u> concede in period t+1 conditional on reaching period t+1. Finally, for any even period t, let $\Pi_1(t,\alpha_2)$ be the expected payoff to player 1 from waiting until period t

****

[8] If $\gamma_i = 1$ for some player i, only a slight modification of the proofs is required. If $\gamma_i = 0$ for both players i, then the set of extended equilibria correspond to the set of all Nash equilibria. If $\gamma_i = 0$ for only one player i, the main results of the paper are still satisfied, but the argument is a bit more complicated due to the possibility that one of the players can move with probability 1 in finite time.

and then conceding, given that the rational player 2 plays strategy α_2. Similarly, $\Pi_1(\infty,\alpha_2)$ is the payoff to player 1 if he plans never to concede, given α_2. For odd periods t, $\mu_2(t-1)$, $\beta_2(t)$, $\Pi_2(t,\alpha_1)$ may be defined similarly.

In what follows, we will assume that both players are following equilibrium strategies. We proceed by establishing some restrictions on the equilibrium strategies.

Lemma 1: Suppose $t > 0$.

(a) If $\beta_i(t) > p_i$, then (i) $\alpha_j(t-1) = 0$ (which implies $\mu_j(t) = 1$) or (ii) $\alpha_j(t+1) = 1$.

(b) $\beta_i(t) < p_i$ implies $\alpha_j(t-1) = 1$.

(c) $\beta_i(t+k) > p_i$ for all even $k \geq 0$ implies $\alpha_j(t-1) = 0$.

Proof:

If, conditional on reaching period t, player i plans to wait with probability greater than p_i, then, conditional on reaching period t-1, player j prefers moving immediately to waiting until period t+1 to move. Since $\gamma_j > 0$ implies that each period is reached with positive probability, it follows that $\Pi_j(t-1,\alpha_i) > \Pi_j(t+1,\alpha_i)$. Therefore, either $\alpha_j(t-1) = 0$, in which case $\mu_j(t) = 1$, or $\alpha_j(t+1) = 1$. This establishes (a).

To establish part (b), note that, if, conditional on reaching period t, player i plans to wait with probability smaller than p_i, then, conditional on reaching period t-1, player j prefers to wait until period t+1 to moving immediately. Therefore, $\Pi_j(t-1,\alpha_i) < \Pi_j(t+1,\alpha_i)$ and hence $\alpha_j(t-1) = 1$.

To establish part (c), recall from the proof of part (a) that

$\beta_i(t+k) > p_i$ implies $\Pi_j(t+k-1,\alpha_i) > \Pi_j(t+k+1,\alpha_i)$. Therefore, if $\beta_i(t+k) > p_i$ for all even $k \geq 0$, then

$$\Pi_j(t-1,\alpha_i) > \lim_{k\to\infty}\Pi_j(t+2k,\alpha_i) = \Pi_j(\infty,\alpha_i)$$

which implies that $\alpha_j(t-1) = 0$ is a best response. Q.E.D.

For $i = 1,2$, define

$$\bar{\mu}_i = (1-p_i)/(1-\gamma_i).$$

Then, using the definitions of $\beta_i(t+1)$ and $\bar{\mu}_i$, we may express $\beta_i(t+1)$ as

$$\beta_i(t+1) = p_i + [1-\mu_i(t)][\alpha_i(t+1)-1] + (1-p_i)[\bar{\mu}_i-\mu_i(t)]/\bar{\mu}_i.$$

Therefore, if $\mu_i(t) > \bar{\mu}_i$, then $\beta_i(t+1) < p_i$ for any value of $\alpha_i(t+1)$ between 0 and 1. It then follows from Lemma 1 that $\alpha_j(t) = 1$. If $\mu_i(t) < \bar{\mu}_i$, then $\alpha_i(t+1) = 1$ implies $\beta_i(t+1) > p_i$. Note, also, that $\bar{\mu}_i < 1$ if and only if $\gamma_i < p_i$.

<u>Lemma 2</u>: $\beta_i(t) > p_i$ implies $\mu_i(t+1) < \bar{\mu}_i$.

<u>Proof</u>:

If $\gamma_i > p_i$, then $\mu_i(t+1) \leq 1 < \bar{\mu}_i$.

If $\gamma_i \leq p_i$, then $\beta_i(t) > p_i$ implies that $\mu_i(t-1) < \bar{\mu}_i$.

Therefore,

$$\mu_i(t+1) = \gamma_i\mu_i(t-1)/\beta_i(t) < \mu_i(t-1) < \bar{\mu}_i .$$

Q.E.D.

Lemma 3: Suppose $t > 1$. If $\mu_j(t-2) < \bar{\mu}_j$ and $1 > K > \mu_i(t-1) \geq \bar{\mu}_i$, then

(a) $\alpha_i(t) = 1$;

(b) $\mu_i(t+1) < \mu_i(t-1)[\gamma_i/[1-K+K\gamma_i]]$.

Proof:

Suppose $\alpha_i(t) < 1$. Then

$$\beta_i(t) = [1-\mu_i(t-1)]\alpha_i(t) + \gamma_i\mu_i(t-1) < 1-(1-\gamma_i)\mu_i(t-1) \leq 1-(1-\gamma_i)\bar{\mu}_i$$
$$= p_i .$$

Lemma 1(b) then implies that $\alpha_j(t-1) = 1$. Our assumption that $\mu_j(t-2) < \bar{\mu}_j$ then implies that $\beta_j(t-1) > p_j$, and from Lemma 1(a) either $\alpha_i(t) = 1$ or $\alpha_i(t-2) = 0$ which implies that $\mu_i(t-1) = 1$. A contradiction.

To establish (b), note that since $\alpha_i(t) = 1$, it follows that

$$\mu_i(t+1) = \mu_i(t-1)[\gamma_i/\beta_i(t)] = \mu_i(t-1)[\gamma_i/[(1-\mu_i(t-1))\alpha_i(t)+\gamma_i\mu_i(t-1)]]$$
$$< \mu_i(t-1)[\gamma_i/[1-K+\gamma_i K]]$$

Q.E.D.

Lemma 4: (a) Suppose $t > 1$. Then $\mu_j(t-2) < \bar{\mu}_j$ and $\mu_i(t-1) \leq \bar{\mu}_i$ implies $\beta_i(t) \geq p_i$.

(b) Suppose $t > 0$. Then $\mu_j(t) < \bar{\mu}_j$ implies either $\beta_i(t) \le p_i$ or $\mu_i(t+1) = 1$.

Proof:

Suppose there is a $t > 1$ such that $\mu_j(t-2) < \bar{\mu}_j$ and $\beta_i(t) < p_i$. Then Lemma 1 implies that $\alpha_j(t-1) = 1$ and hence that $\beta_j(t-1) > p_j$. But Lemma 1 then implies that either $\alpha_i(t) = 1$ or $\mu_i(t-1) = 1$. If we now suppose that $\mu_i(t-1) \le \bar{\mu}_i$, then it follows by definition $\beta_i(t) \ge p_i$. This contradiction establishes part (a).

To establish part (b), suppose there is a $t > 0$ such that $\mu_j(t) < \bar{\mu}_j$ and $\beta_i(t) > p_i$. Then Lemma 1(a) implies that $\alpha_j(t+1) = 1$ and therefore that

$$\beta_j(t+1) = \alpha_j(t+1)[1-\mu_j(t)] + \mu_j(t)\gamma_j = [1-\mu_j(t)] + \mu_j(t)\gamma_j$$
$$= 1 - \mu_j(t)[1-\gamma_j] > 1-\bar{\mu}_j+\bar{\mu}_j\gamma_j = p_j.$$

It also follows from Lemma 2 that $\mu_i(t+1) < \bar{\mu}_i$. Proceeding by induction, we may establish that $\beta_j(t+k) > p_j$ for all odd $k > 0$. It then follows from Lemma 1(c) that $\alpha_i(t) = 0$ which implies that $\mu_i(t+1) = 1$. Q.E.D.

Lemma 5: If $\gamma_2/p_2 > 1 \ge \gamma_1/p_1$, then $\mu_1(1) = 1$.

Proof:

Suppose $\mu_1(1) < 1$. Note first that $\gamma_2 > p_2$ implies that $\bar{\mu}_2 > 1$. Therefore, if $\mu_1(1) \ge \bar{\mu}_1$, then it follows by induction on Lemma 3(b) that $\mu_1(t) < \bar{\mu}_1$ for some $t < \infty$. Let $\hat{t}$ be the smallest such t. Then Lemma 4(a) implies that $\beta_1(\hat{t}+1) \ge p_1$ from which it follows that

$$\mu_1(\hat{t}+2) \leq \mu_1(\hat{t})\gamma_1/p_1 \leq \mu_1(\hat{t}) < \bar{\mu}_1 \leq 1.$$

Letting $i = 1$ in both parts (a) and (b) of Lemma 4, it follows that $\beta_1(\hat{t}+1) = p_1$. Proceeding by induction, we may conclude that $\beta_1(t+1) = p_1$ and $\mu_1(t) < \bar{\mu}_1 \leq 1$ for odd $t > \hat{t}$.

Now consider player 2. For $t \geq \hat{t}+2$, Lemma 4 implies that either $\beta_2(t) = p_2$ or $\mu_2(t+1) = 1$. Therefore, if $\mu_2(t+1) < 1$, it follows by definition that

$$\mu_2(t+1)/\mu_2(t-1) = \gamma_2/p_2 > 1.$$

By induction, we may conclude that $\mu_2(\tilde{t}) = 1$ for some $\tilde{t} > \hat{t}$. But this implies that $\beta_2(t) = \gamma_2 > p_2$ for all odd $t > \tilde{t}$. Then, by Lemma 1(c), $\alpha_1(\tilde{t}) = 0$ and hence $\mu_1(\tilde{t}+1) = 1$. This contradiction proves the result. Q.E.D.

Define $\hat{t}_i = \sup\{t: \mu_i(t) < 1\}$.

<u>Lemma 6</u>: Suppose $\gamma_i/p_i > 1$ for $i = 1,2$. Then (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) $|\hat{t}_2-\hat{t}_1| = 1$.

<u>Proof</u>:

Since $\gamma_i/p_i > 1$ implies $\bar{\mu}_i > 1$, part (a) follows immediately from Lemma 4.

To establish (b), suppose that $\hat{t} = \infty$. Then since $\bar{\mu}_i > 1$ for

$i = 1,2$, it follows from Lemma 4 that $\beta_i(t) = p_i$ and hence that $\mu_i(t) = (\gamma_i/p_i)\mu_i(t-2)$ for all $t > 1$. But since $\gamma_i/p_i > 1$, this 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 $k > 0$, $\beta_i(k+t) = \gamma_i > p_i$. Then Lemma 1(c) then implies that $\alpha_j(t) = 0$ and hence $\mu_j(t+1) = 1$. Q.E.D.

<u>Lemma 7</u>: Suppose $\gamma_i/p_i > 1$, $i = 1,2$. Then

(a) $\hat{t}_1 = \hat{t}_2-1$ implies

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

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

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

<u>Proof</u>:

Since $\gamma_i/p_i > 1$ implies that $\bar{\mu}_i > 1$, it follows from Lemma 4(a) that $\beta_i(\hat{t}_i+1) \geq p_i$. Suppose $\hat{t}_i = \hat{t}_j-1$. If $\beta_i(\hat{t}_i+1) > p_i$, then by Lemma 1(a), either (i) $\mu_j(\hat{t}_j) = \mu_j(\hat{t}_i+1) = 1$ or (ii) $\alpha_j(\hat{t}_i+2) = 1$ which implies that $1 = \mu_j(\hat{t}_i+3) \leq \mu_j(\hat{t}_i+1) < 1$. In either case, we have a contradiction, and, therefore, $\beta_i(\hat{t}_i+1) = p_i$. Consequently, $\hat{t}_1 = \hat{t}_2-1$ implies

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

Moreover, we have already established that $\beta_2(\hat{t}_2+1) \geq p_2$. Therefore,

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

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

In order to use Lemmata 1 through 7 to establish our main results, we must first verify that a subgame perfect equilibrium exists for $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ for all small ϵ.

Lemma 8: Suppose $\gamma_2 > p_2$ and $\gamma_2/p_2 \geq \gamma_1/p_1$. Then, for any game $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ with $0 < \epsilon < p_i/\gamma_i$, $i = 1,2$, a subgame perfect equilibrium exists.

Proof:

Suppose first $\gamma_1/p_1 \leq 1$. Let $\alpha_1 = (0,0,0,\ldots)$ and $\alpha_2 = (1,1,1,\ldots)$. Since $\alpha_1(0) = 0$, it follows that $\mu_1(t) = 1$ and hence $\beta_1(t+1) = \gamma_1 \leq p_1$ for all odd $t > 0$. It may then be verified that, for all odd t, $\alpha_2(t) = 1$ is a best response. Furthermore, since $\bar{\mu}_2 > 1$, $\alpha_2(t) = 1$ implies that $\beta_2(t) > \gamma_2 > p_2$ for all even $t > 0$. It may be readily verified that $\alpha_1 = (0,0,0,\ldots)$ is the best response.

Suppose next that $\gamma_1/p_1 > 1$. Define $\hat{t}$ to be the unique even integer t which satisfies

$$\epsilon(\gamma_2/p_2)^{t/2} < 1 \leq \epsilon(\gamma_2/p_2)^{(t+2)/2}.$$

Then define $\alpha_1(0)$ so that $\epsilon\gamma_1+(1-\epsilon)\alpha_1(0) = \beta_1(0) = \gamma_1\epsilon(\gamma_1/p_1)^{\hat{t}/2}$ and choose α_1 so that $\beta_1(t) = p_1$ for even $t = 2,4,\ldots,\hat{t}$, and $\alpha_1(t) = 0$ for even $t > \hat{t}$. Then it may be verified that $\mu_1(\hat{t}+1) = 1$ and hence that $\beta_1(t) = \gamma_1 \geq p_1$ for even $t > \hat{t}$. Choose α_2 so that $\beta_2(t) = p_2$ for odd $t < \hat{t}$ and $\alpha_2(t) = 0$ for odd $t \geq \hat{t}+1$. Then $\mu_1(\hat{t}+2) = 1$ and hence $\beta_2(t) = \gamma_2 > p_2$ for odd $t > \hat{t}+1$. The value of $\hat{t}$ has been chosen so that $\mu_2(\hat{t}) = \epsilon(\gamma_2/p_2)^{\hat{t}/2} \geq p_2/\gamma_2$. Therefore, $\beta_2(\hat{t}+1) = \mu_2(\hat{t})\gamma_2 \geq p_2$. It may be verified that these strategies form a pair of best responses. Q.E.D.

The next Lemma establishes some properties of the subgame perfect equilibrium for $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ as ϵ becomes small.

Lemma 9: Suppose $\gamma_2/p_2 \geq \gamma_1/p_1 > 1$. Then for any $t > 0$ and any $\epsilon_1 > 0$, there is a $\psi > 0$ such that, for all $\epsilon < \psi$: (a) $\beta_2(1) = p_2$, (b) $\hat{t}_i > t$, and (c) $|\alpha_i(t)-p_i| < \epsilon_1$.

Proof:

We establish first that $\beta_2(1) = p_2$ for $\epsilon > 0$ sufficiently small. Suppose $\beta_2(1) < p_2$. Then

$$\mu_2(2) = \mu_2(0)\gamma_2/\beta_2(1) > \mu_2(0)\gamma_2/p_2 = \epsilon\gamma_2/p_2.$$

Furthermore, Lemma 1 implies that $\alpha_1(0) = 1$ and hence $\mu_1(1) < \mu_1(-1) = \epsilon$. It then follows from Lemma 7 (a and b) that

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

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 $\mu_2(2) \leq \epsilon\gamma_2/p_2 < 1 < \bar{\mu}_2$ for ϵ sufficiently small. Then, since $\mu_1(1) \leq 1 < \bar{\mu}_1$, part (a) follows from Lemma 4a.

To establish part (b), note that Lemma 7 then implies that

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

Therefore, for any $t > 1$, there is an $\psi > 0$ such that $\epsilon < \psi$ implies $t < \hat{t}_2$. We may conclude, therefore, that $\hat{t}_2 \to \infty$ as $\epsilon \to 0$.

Combined with part (b) Lemma 6 then implies that for $\epsilon < \psi$, $\beta_1(k) = p_1$ for even k, $0 < k \leq t$. Lemma 7 then yields

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

Since $\hat{t}_1 \to \infty$ as $\epsilon \to 0$, it follows that $\mu_1(t) \to 0$ as $\epsilon \to 0$. Then since $\beta_1(t) = p_1$, it follows from the definition of $\beta_1(t)$ that $\alpha_1(t) \to p_1$ as $\epsilon \to 0$.

A similar argument establishes that, for any odd $t > 0$, $\alpha_2(t) \to p_2$ as $\epsilon \to 0$. Q.E.D.

Proof of Theorems 1 and 2:

We establish first Theorem 2a. Suppose $\gamma_2/p_2 > \gamma_1/p_1 > 1$. Then the definitions of $\mu_1(1)$ and $\beta_1(0)$ combined with Lemma 7 (a and b) imply that

$$[\epsilon\gamma_1+(1-\epsilon)\alpha_1(0)]/\gamma_1 = \beta_1(0)/\gamma_1 = \epsilon/\mu_1(1) = (p_2/\gamma_2)\mu_2(2)/\mu_1(1)$$
$$\leq [\gamma_1 p_2/\gamma_2 p_1]^{(\hat{t}_2+2)/2}.$$

Then since Lemma 9 implies that $\hat{t}_2 \to \infty$ as $\epsilon \to 0$, it follows immediately that $\alpha_1(0) \to 0$ as $\epsilon \to 0$. Lemma 9 also implies that $\alpha_i(t) \to p_i$. Theorem 2a then follows from Lemma 8 and the definition of the metric d.

To prove Theorem 2(b), we note that if $\gamma_2/p_2 > 1 > \gamma_1/p_1$, then since Lemma 5 implies $\alpha_1(0) = 0$, it follows that $\mu_1(t) = 1$ for all odd $t > 0$. Therefore, $\beta_1(t) = \gamma_1 < p_1$ for even $t > 0$. It then follows from Lemma 1(b) that, for all odd t, $\alpha_2(t) = 1$ is the unique best response and hence $\beta_2(t) > \gamma_2 > p_2$. Lemma 1(c) then implies that, for all even t, $\alpha_1(t) = 0$ is the unique best response. Theorem 2(b) then follows from Lemma 8 and the definition of the metric d.

Theorem 1 follows from Theorem 2 except for the case where $\gamma_2/p_2 > 1 = \gamma_1/p_1$. But, in this case, Lemma 5 implies that $\alpha_1(0) = 0$. The theorem then follows from Lemma 8 and the definition of the metric d. Q.E.D.

Proof of Theorem 3(a):

Suppose that $\gamma_2/p_2 = \gamma_1/p_1 > 1$. To establish that $p_1 \leq \alpha_1(0) \leq \gamma_1$, note first that Lemma 9 implies that $\beta_2(1) = p_2$ and hence $\mu_2(2) = (\gamma_2/p_2)\epsilon$. From Lemma 7, combined with Lemma 6, we know that $\epsilon \leq \mu_1(1) \leq \mu_2(2)$. Therefore,

$$\epsilon \leq \mu_1(1) = \epsilon(\gamma_1/\beta_1(0)) = \epsilon\gamma_1/[(1-\epsilon)\alpha_1(0)+\epsilon\gamma_1] \leq \mu_2(2) = \epsilon(\gamma_1/p_1).$$

Therefore, if a subsequence of $\alpha_1(0)$ converges as $\epsilon \to 0$, then, in the limit,

$$p_1 \leq \alpha_1(0) \leq \gamma_1.$$

To complete the proof of Theorem 3(a), it is sufficient to show that for any $\bar{x} \in [p_1,\gamma_1]$, there is a sequence $\{\epsilon_t\}$ such that (i) $\epsilon_t \to 0$ and (ii) $\alpha_1(0) = \bar{x}$ for some subgame perfect equilibrium of $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon_t)$. For each positive integer t such that $(\gamma_1/p_1)^{t/2} > 1$, let $\hat{\epsilon}_t \in (0,1)$ be the unique solution to

$$\epsilon\gamma_1 + (1-\epsilon)\bar{x} = \gamma_1\epsilon(\gamma_1/p_1)^{t/2}.$$

Note that $\epsilon_t \to 0$ as $t \to \infty$.

Now consider the sequence of games $\{(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon_t)\}$. For each even t, choose α_1 so that $\alpha_1(0) = \bar{x}$, $\beta_1(k) = p_1$ for even k, $0 < k \leq t$, and $\alpha_1(k) = 0$ for even $k > t$. Note that $\mu_1(t+1) = 1$ and hence $\beta_1(k) = \gamma_1 > p_1$ for all even $k > t$. Define α_2 so that $\beta_2(k) = p_2$ for odd $k < t$ (This is possible since $\gamma_2/p_2 = \gamma_1/p_1$ so that for all even $k < t$, $\mu_2(k-1) \leq p_2/\gamma_2$.) and $\alpha_2(k) = 0$ for all odd $k > t$. Note that $\gamma_2/p_2 = \gamma_1/p_1$ also implies $\beta_2(t+1) \geq p_2$. Therefore, α_1 and α_2 form a pair of best responses. Q.E.D.

<u>Proof of Theorem 3(b)</u>:

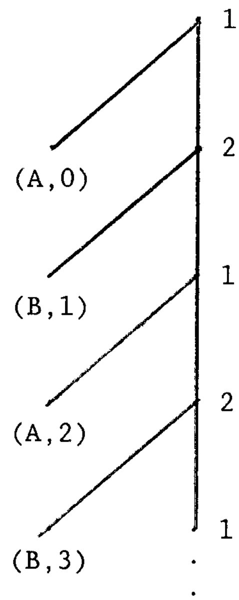
It is enough to verify that if $\gamma_2/p_2,\gamma_1/p_1 \leq 1$, then

(i) $\alpha_1 = (0,0,0,\ldots)$ and $\alpha_2 = (1,1,1,\ldots)$ and (ii) $\alpha_1 = (1,1,1,\ldots)$ and $\alpha_2 = (0,0,0,\ldots)$ are both extended subgame perfect equilibria.

Consider the game $(\Gamma(\delta_1,\delta_2),(\gamma_1,\gamma_2),\epsilon)$ with $0 < \epsilon < \bar{\mu}_i$, $i = 1,2$. Suppose $\alpha_2 = (1,1,1,\ldots)$. Then, for all even $t > 0$, $\mu_2(t) < \mu_2(t-2) < \ldots < \mu_2(2) < \epsilon < \bar{\mu}_2$. Therefore, $\beta_2(t) > p_2$ for all odd $t > 0$. Lemma 1(c) then implies that, for all even $t \geq 0$, $\alpha_1(t) = 0$ is a best response. On the other hand, if $\alpha_1 = (0,0,0,\ldots)$, then $\mu_1(t) = 1$ for all odd $t > 0$ and hence $\beta_1(t) = \gamma_1 \leq p_1$ for all even $t > 0$. Therefore, for all odd $t > 0$, $\alpha_2(t) = 1$ is a best response. This establishes case (i).

The argument for case (ii) is similar. Q.E.D.

Figure 1. The Extensive Game



References

Binmore, K. (1985),"Modeling Rational Players," CARESS Working Paper, #85-36, University of Pennsylvania, December.

Fudenberg, D. and E. Maskin (1986), "The Folk Theorem for Repeated Games with Discounting or with Incomplete Information," forthcoming in Econometrica, 54 (1986).

Hendricks, K. and C. Wilson (1985), "The War of Attrition in Discrete Time," C.V.Starr Center Economic Research Report 85-32, September.

Kohlberg, E. and J.F. Mertens (1985), "On the Strategic Stability of Equilibria," Harvard Business School Working Paper 1-785-012, November.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982), "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," Journal of Economic Theory, 27, pp.245-252.

Kreps, D. and R. Wilson (1982), "Reputation and Imperfect Competiton," Journal of Economic Theory, 27, pp.253-279.

Meyerson, R. (1978), "Refinements of the Nash Equilibrium Concept," International Journal of Game Theory, 4, pp.73-80.

Milgrom, P. and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence," Journal of Economic Theory, 27, pp.280-312.

Selten, R. (1975), "A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games,", International Journal of Game Theory 4, 25-55.

Wilson, R. (1985), "Reputations in Games and Markets," in Game Theoretic Models of Bargaining, edited by A. Roth, Cambridge University Press.