Discussion Paper No. 896

A COMMUNICATION-PROOF EQUILIBRIUM CONCEPT

by

José Luis Ferreira*

Northwestern University

August 1990

* The author acknowledges helpful comments from P. DeMarzo, A. Manelli and, especially, from E. Kalai, whose support has made this work possible. Of course, any error is the author's.

A COMMUNICATION-PROOF EQUILIBRIUM CONCEPT

by

Jose-Luis Ferreira

Northwestern University

August 1990

0. ABSTRACT

This paper proposes an equilibrium concept for the classes of environments in which players can communicate with each other but cannot make binding agreements. This Communication-proof equilibrium is intended to be regarded as an extension of both Coalition and Renegotiation-proof equilibria. Conceptual foundations for this particular definition are widely discussed as it is confronted with other definitions in this class of environments. The definition is extended to infinite horizon games using the von Neumann and Morgenstern's concept of abstract stable sets.

1. INTRODUCTION

Traditionally, games have been divided in two classes, cooperative and non cooperative. What defines cooperation is the players' capability to communicate and make binding commitments, whereas the lack of these abilities leads to the non cooperative scenarios. But, of course, two classes of games cannot exhaust the possibilities of a division that is made upon two different characteristics.

Recently, many works have been dedicated to the class of environments in which players can freely communicate with each other, but cannot make binding agreements. The Coalition-Proof Nash equilibrium (CPNE) and different Renegotiation-Proof equilibria (RPE) have been defined as reasonable solution concepts for some particular situations.

Bernheim, Peleg and Whinston (1986) (from now on B,P&W) defined the

CPNE for normal form games in the spirit of the Nash equilibrium; since all players move simultaneously, they allow for a deviating coalition that can take as given the opponent coalition's strategy; but when a first coalition considers whether to deviate, it should know that a subcoalition may consider further deviations, and so on[1]. Despite the fact that it is believed that the CPNE may not recognize all possible deviations that can "credibly" occur, it has been widely accepted as a "consistent" attempt to describe coalitional behavior in games of simultaneous moves.

In multi-stage games, the definition of a RPE has been studied in several works. The idea is that players will not submit to a "grim" strategy in a subgame if they can renegotiate to a better equilibrium for all of them. For finite horizon games, backwards induction allows for a natural definition of RPE, the Pareto Perfect equilibrium, given by Bernheim and Ray (1989) (B&R). Farrel and Maskin (1989) (F&M), Bernheim and Ray (1989) and Asheim (1988) extended this concept to infinite horizon games in different ways; but no one of them has provided a generally accepted definition.

In all the literature of renegotiation-proof, only the coalition of all players can renegotiate. However, as pointed out by F&M, "a potential improvement by all players needs not be a prerequisite to renegotiation; a proper subset may profit from renegotiating by themselves."

B,P&W extended the CPNE to games in extensive form (the Perfectly Coalition-Proof Nash equilibrium, PCPNE); but, as it will be argued, it does not fully capture the idea of renegotiation.

---

[1] The earlier definition of Strong Nash equilibrium in Aumann (1959) allows for coalitional deviations even if they are not immune to further deviations.

The purpose of this paper is to define an equilibrium that may be regarded as both coalition and renegotiation proof; the Communication-Proof equilibrium (Com-PE). Following recommendations in Abreu and Pearce (1989), conceptual foundations for this particular definition will be discussed (in the sense of being explicit about the communication process and the way in which players renegotiate and deviate).

To better understand the discussion, let us consider the example in figure 1 (this example is in Peleg (1988)):

FIGURE 1

It is straightforward to check that $A=(l_1, r_2, l^*_1, l_3)$ and $B=(l_1, l_2, r^*_1, r_3)$ are the only Subgame Perfect equilibria of this game in pure strategies. If, for whatever reason, B is proposed, the three players have a clear incentive to renegotiate to A at the subgame in which player 2 moves. No similar deviation by the grand coalition can be found in A. This renegotiation process isolates A as the only PPE in this game.

It can be shown (Peleg 1988) that this game has no PCPNE. According to definition 5 below, it has to be checked that the proposed strategy profile is a PCPNE in all subgames and in all games induced by the strategy of any subset of players. B cannot be a PCPNE by the argument above and because (r_2, l^*_1, l_3) is a PCPNE in the game induced by r_1. To show that A is not a PCPNE, consider the game induced by r_2 (figure 2):

FIGURE 2

The coalition formed by players 1 and 3 will clearly "renegotiate" from (l_1, l^*_1, l_3), the strategy profile induced by A in this game, to (r_1, r^*_1, r_3), since this is a PCPNE in this induced game, it is enough to rule out A as a PCPNE.

A main objection can be pointed out at this moment: when considering a deviation, taking as given the opponents' strategies is a natural assumption only if those strategies and the deviation are played simultaneously. If part of the deviation can be observed by some opponent, the possibility of a reaction by the opponents cannot be ruled out. In the example, after r_1, player 2 observes the deviation and should be able to prevent the path (r_2, r^*_1, r_3) that, after all, is not an equilibrium. The precise way in which this is done will be described below, after the definition of the Communication-proof equilibrium is given.

2. NOTATION

Γ: a game.

$N = \{1, \ldots\}$, the set of players. A subset of N will denote a coalition

S_i will denote the (compact) set of strategies for player i; $S = X_{i\in N} S_i$;

$S_C = X_{i\in C} S_i$ and $S_{-C} = X_{i\notin C} S_i$, where $C\subset N$ is a coalition and $-C=N\backslash C$.

Their respective typical elements will be s_i, s, s_C and s_{-C}.

$u^i : S\to R$ is the (continuous) outcome function for player i.

$\Gamma|s_C$ is the game that s_C induces on Γ. For details see Peleg (1988).

g will denote a subgame of Γ.

$s(g)$ is the strategy induced by s in the subgame g.

H is the set of feasible histories of Γ; $h \in H$ is a history.

H^h is the set of feasible histories following h.

g^h is the subgame induced by history h.

$s|_h$ is the strategy induced by s after history h.

$s(t)$ is the behavior that strategy s induces at stage t.

$t(h)$ stage at which history h is observed (history h ends at stage $t(h)-1$).

$h(1)$ set of histories of length $t(h)+1$ that belong to H^h.

The number of stages for a given game is the maximum number of nested subgames in it.

For details regarding the definition of feasible histories, see Asheim (1988).

When there is no confusion, the short name of an equilibrium (e.g SPE) will denote the set of those equilibria in a given game (that set will also be denoted by e.g SPE(Γ)).

3. CONCEPTS

In what follows, games will be assumed to be of complete information and perfect recall so that only behavioral strategies will be considered.

DEFINITION 1. s^* is a <u>Nash equilibrium restricted to $T \subset S$</u>, NE(T), iff $s^* \in T$ and for all $i \in N$ and all $s_i \in S_i$ such that $(s^*_{-i}, s_i) \in T$, $u^i(s^*_{-i}, s_i) \leq u^i(s^*)$.

When $T=S$, we have the standard definition of a Nash equilibrium (NE).

DEFINITION 2. (i) In a single player game, s^* is a Coalition-proof Nash equilibrium restricted to $T \subset S$, CPNE(T), iff $s^* \in \text{argmax}_{s \in T}\, u^1(s)$.

(ii) Assume that CPNE(T) has been defined for games with less than n players. Then, in a n-players game;

(a) $s^* \in S$ is Self-enforcing restricted to T, SE(T), iff for any coalition $C \neq N$, s^* is CPNE(T) in the game $\Gamma | s^*_{-C}$;

(b) $s^* \in S$ is CPNE(T) if it is SE(T) and if it does not exist any other $s \in T$ such that s is SE(T) and $u^i(s) > u^i(s^*)$ for all $i \in N$.

If T=S, CPNE(T) is the definition of coalition-proof Nash equilibrium (CPNE) given by B,P&W.

DEFINITION 3. $s^* \in S$ is a Subgame perfect equilibrium, SPE, iff, for any subgame g, $s^*(g)$ is a NE.

If Γ has a finite number of stages, say t, an alternative definition of SPE, using definition 1 above, is as follows:

DEFINITION 3′. Let Γ be an extensive form game with $t(<\infty)$ stages.

(i) If $t=1$, $s^* \in S$ is a SPE′ iff s^* is a NE.

(ii) Assume that a SPE′ has been defined for all games with $r<t$ stages and consider a game with t stages; then $s^* \in S$ is SPE′ iff s^* is a NE(T) where $T = \{s \in S \,/\, s \text{ induces a SPE′ in proper subgames of } \Gamma\}$.

PROPOSITION 1. s^* is a SPE if and only if it is a SPE′.

PROOF. See appendix.

DEFINITION 4. (i) In a single stage game Γ, s^* is a <u>Communication proof equilibrium</u>, Com-PE, iff it is a CPNE.

(ii) Let $t>1$ and assume that Com-PE has been defined for games with $r<t$ stages. Then, in a game Γ with t stages, s^* is <u>Com-PE</u> iff it is a CPNE(S^1) where $S^1=\{s\in S$ / s induces a Com-PE in proper subgames of $\Gamma\}$.

REMARK 1. If we replace CPNE with NE in definition 4, we obtain SPE and Com-PE is seen as a natural extension of definition 3'.

REMARK 2. s is a Com-PE iff for every history h and every coalition C, $(s_C(t(h)), s(h(1)))$ is a Com-PE in $g^h|s_{-C}(t(h))$. i.e. at the beginning of any subgame, and when the strategies by any coalition are fixed for the current period, s is a Com-PE.

REMARK 3. The name CPNE is not necessary in definition 4; for one-stage games, the Com-PE can be defined as the CPNE and then the recursive definition may continue. If a new definition of coalition-proof equilibrium is presented for one-stage games, a new definition of communication-proof is immediately available.

The following proposition studies the existence of Com-PEa for a special and important class of games.

PROPOSITION 2. Let Γ be a <u>perfect information finite game</u> (the number of stages, players and strategies is finite); then there exists a Com-PE.

PROOF. Perfect information means that only one player moves at each stage of the game. The proof is inductive in the number of stages, t.

Let t=1, then a Com-PE exists by the finiteness of alternatives and transitivity of preferences.

Let t>1 and assume the proposition is true for games of r<t stages. W.l.o.g. let player one be the (only) player moving at the first stage in Γ. Consider the (finite) set of Com-PEa in each subgame $g(k_i)$, where k_i is a one-stage history. Given $g(k_i)$, select a Com-PE that maximizes player one's utility and denote it by $s_1(k_i)$; (such a Com-PE exists by finiteness of Com-PEa and by the transitivity of the preferences). Once $s_1(k_i)$ is selected for each of the m one-stage histories take $s_1 = (s_1(k_1), \ldots, s_m(k_m))$. Let $s^1(s_1)$ be a strategy that maximizes player one's expected utility at stage one when s_1 is to be played thereafter. Then $s^* = (s^1(s_1), s_1)$ is a Com-PE. If not, there exist a coalition $C \subset N$ and a strategy $s'_C \in S_C$ such that $s' = (s'_C, s^*_{-C})$ is a CPNE(S') in the game $\Gamma | s^*_{-C}(1)$ and $s' >_C s^*$. Since player one cannot improve his utility in any subgame (by construction), he is not in C; but then $s_1(k_i)$ is not a Com-PE for some k_i played with positive probability in $s^1(s_1)$.

Proposition 3 relates the concept of Com-PE with those of Pareto Perfect equilibrium and Perfectly Coalition-Proof Nash equlibrium, which are defined below for the sake of completeness.

DEFINITION 5. (B,P&W). (i) In a single player, single stage game Γ, $s^*\in S$ is a <u>Perfectly Coalition-Proof Nash equilibrium</u> (PCPNE) iff s^* maximizes $u^i(s)$.

(ii) Let $(n,t)\neq(1,1)$. Assume that PCPNE has been defined for all games with m players and s stages, where $(m,s)\leq(n,t)$ and $(m,s)\neq(n,t)$.

a) For any game Γ with n players and t stages, $s^*\in S$ is <u>Perfectly self-enforcing</u> (PSE) if (1) for all $C\subset N$, s^*_C is a PCPNE in the game $\Gamma|s^*_{-C}$, and (2) for any proper subgame of Γ, g, $s^*(g)$ is a PCPNE in g.

b) For any game Γ with n players and t stages, $s^*\in S$ is <u>PCPNE</u> if it is PSE and if there does not exist another PSE strategy vector $s\in S$ such that $u^i(s)>u^i(s^*)$ for all $i\in N$.

REMARK 4. As pointed out by B,P&W, being PCPNE is not equivalent to being CPNE in every subgame (a possible extension of definition 3 to coalitions).

DEFINITION 6. (B&R). (i) In a single stage game, a <u>Pareto-perfect equilibrium</u> (PPE) is a Nash equilibrium that is not strictly Pareto-dominated by another Nash equilibrium.

(ii) Let $t>1$ and assume that PPE has been defined for all games with less than t stages. Let Γ be a t stages game; then a strategy profile s is a <u>PPE</u> in Γ iff

(a) s is a Nash equilibrium, and $s|_h$ is a PPE of g^h for all $h\in H$, $h\neq\emptyset$, and

(b) there is no profile x satisfying part (a) such that $u_i(s)<u_i(x)$

for all $i \in N$.

PROPOSITION 3. Let Γ be a finite horizon game, then

(i) Com-PE $\subset$ SPE.

(ii) If Γ is a one-stage game ($t=1$), Com-PE = CPNE.

(iii) If Γ is a two-player game ($n=2$), Com-PE = PPE.

(iv) Neither Com-PE $\subset$ PCPNE nor PCPNE $\subset$ Com-PE is satisfied.

PROOF.

(i) Follows from proposition 1 and from the fact that CPNE $\subset$ NE.

(ii) Follows from the definition of Com-PE.

(iii) The proof procedes by induction on the number of stages of the game. For $t=1$ the statement is trivial (CPNE reduces to a non strictly Pareto-dominated Nash equlibrium). Assume now that the proposition has been proven for $s<t$, and let Γ be a t-stages game. First prove PPE $\subset$ Com-PE; let $s \in S$ be a PPE, by definition 6 (ii)(a) $s \in S^* = \{s \in S$ / s induces a PPE in proper subgames$\}$, by induction hypothesis $S^* = S' = \{s \in S$ / s induces a Com-PE in proper subgames$\}$; also $s \in NE$, what, for $n=2$, means that s is self-enforcing; these two facts and definition 6 (ii)(b) show that s is CPNE restricted to S'. To prove Com-PE $\subset$ PPE, let s be CPNE restricted to S'; $s \in S'$ implies that $s \in S^*$ by induction hypothesis. By definition of Com-PE we also have that

(*) it does not exist any other $s' \in S'$ such that $s_j = s'_j$ and $s' >_j s$ for all $i \in \{1,2\}$ and $j \neq i$.

This means that s is a NE and definition 6 (ii)(a) is satisfied; if not, there exists a player $i \in \{1,2\}$ and a strategy $s'_i \in S_i$ such that

$s'=(s'_i, s_j) >_i s$. But then $s(1)=s'(1)$ (otherwise $s(1)$ was not a NE in subgames) and $s' \in S'$, which contradicts (*). Finally, part (ii)(b) of definition 6 is satisfied by the optimality of CPNE.

(iv) See counterexample in the appendix.

Let Γ be an extensive game form with perfect information. Peleg (1988) shows that, for every profile of linear preferences (i.e when there are no indiferences), the set of SPE of Γ coincides with the set of PCPNE; he also studies some important applications of this result to the theory of voting. Proposition 4 shows that the Com-PE satisfies the same property and, therefore the same applications follow.

PROPOSITION 4. Let Γ be a finite horizon perfect information game with linear preferences. Then the Com-PE(Γ)=SPE(Γ).

PROOF. The proof is straightforward and contrasts with the more involved proof for PCPNE given by Peleg.

For one-stages games ($t=1$), since there are no simultaneous moves, Com-PE=NE=SPE trivially. Also, by the linearity of preferences, the equilibrium is unique.

Assume that Com-PE=SPE for $s<t$, and that the equilibrium is unique. Let Γ be a t-stage game. Denote by $s^*(h_i)$ the only SPE (and CPE) after the one-stage history h_i; let $s^*(1)=(s^*(h_1),\dots,s^*(h_m))$, where m is the number of one-stage histories (i.e. the number of alternatives available to the first player moving -w.l.o.g. player one-). Given $s^*(1)$, at the beginning of the game, player one will choose the alternative (the one-stage

history) that induces his preferred equililibrium afterwards; denote that alternative by h^*_i (h^*_i is unique by linearity of preferences), then $s^*=(h^*_i, s^*(h^*_i))$ is easily seen as the only SPE and CPE of Γ, since no deviation (coalitional or individual) can take place after the first stage (by the uniqueness of CPE thereafter) and any deviation in the first stage will make player one worse off (by construction).

Definitions 1, 2, 3 and 4 are applicable to any extensive form game; but, of course, NE and CPNE make more sense when applied to games in normal form. A normal form game can be thought as an extensive form game (with perfect recall) of one stage in which all possible actions by one player lead to the same information set of some other player (this reflects the fact that they play simmultaneously). When players do not play simmultaneously in the same stage, the refinements of perfect and sequential equilibria rule out some "irrational" moves in unreached information sets. The definition of Com-PE does not deal with these issues. It generalizes the definition of SPE to coalitions in the same spirit as CPNE generalizes that of NE. Possible refinements are postponed to future research.

4. DISCUSSION

In the example in section 1, it was said that, after $A=(l_1,r_2,l^*_1,l_3)$ is proposed, the deviation $D=(r_1,r^*_1,r_3)$ can be objected because it results in the path (r_1,r_2,r^*_1,r_3), that is not an equilibrium at the point when player 2 has to move. To predict the deviation D after A is to predict the strategy profile $P=(r_2,r^*_1,r_3)$ after r_1 as an equilibrium of that subgame; but it is not clear at all that the communication within the coalition (1,3) at the beginning of the game can change the nature of P from being a non-equilibrium path to being an equilibrium one. Instead, it seems better to think that, whenever communication can take place, the continuation of the strategy should be an equilibrium that prevents any deviation that may occur after that precise communication (in the case of no communication, the continuation should, at least, be a SPE). Once the set of equilibria in the last stage has been detected, one stage earlier any coalition C has to consider that any deviation s'_C to a strategy profile s, has to satisfy that (s'_C,s_{-C}) induces an equilibrium in the last stage, because communication (renegotiation) at the beginning of that last stage will lead to an equilibrium.

In the example in section one, one may think of the following procedure: first, A is proposed, then coalition (1,3) considers the deviation D; if D occures, player 2 observes r_1 instead of l_1 and calls for renegotiation, the three players sit at the negotiation table and decide how to play thereafter; the only "good" equilibrium at this point is (r_2,l^*_1,l_3). Knowing this, player one should never agree to deviate from l_1 to r_1 and A remains an equilibrium. In this particular example it is not even necessary that communication takes place; it is enough that players can reproduce the argument. Without communication, it is also possible to think that, after player 2 has observed r_1 (instead of l_1), he may deduce that the only explanation for that is the deviation D; then his best response is to play l_2 (instead of r_2) and this, again, rules out D as a profitable deviation for player one (see section 5 below for an example and a discussion concerning forward induction and communication).

The general scheme to deal with communication issues is proposed as follows:

1) There is a definition of equilibrium E to be applied in a multi-stage game in which players can freely communicate with each other at the beginning of each stage (subgame).

2) At the beginning of the game, a strategy profile $s \in S$ is proposed; this will be called a "social agreement"[2]. s will be also proposed at the beginning of every stage if a deviation has been observed by a player outside the deviating coalition.

3) If a coalition C deviates from s at stage t using $s'_C \in S_C$, this deviation constitutes a new social agreement for members of C.

4) Every player obeys the last social agreement in which he is involved. If, by doing so, the result is an equilibrium E ex-ante (in the sense that, up to the extent of their abilities, beliefs and knowledge, players do not find any incentive to form new social agreements), then that last social agreement will be followed until the end of the game.

5) Every player is aware of all social agreements of every coalition he has been in (including s for the coalition of all players at the beginning of any stage when it is proposed).

6) Any player i believes that any other player j will behave according to the last social agreement involving j that player i is aware of.

---

[2] A social agreement is a unique element of S, unlike a social norm, which usually refers to a subset of S. Later on, a social agreement that is also an equilibrium will be regarded as part of a social norm.

7) If, at stage t, coalition C plans a deviation that consists of a change of strategy at stages $t_{\alpha 1}, t_{\alpha 2}, \ldots, t_{\alpha 3}$, the time at which a given player knows the change at stage t_* is at

i) stage t if he belongs to C and

ii) stage t_*+1 if he does not.

8) No deviation takes place after s has been proposed if and only if s satisfies the definition of equilibrium E.

The concept of Com-PE fits in this scheme. If no communication is permitted (no coalitions are allowed) so does the SPE and, for one-stage games, the NE. The CPNE also fits for one-stage games. The PPE follows the scheme when only the grand coalition is allowed to be formed. However, the PCPNE does not fit in it since players do not react to deviations (by renegotiating to the first agreement, as said in 2)).

Consider now the following situation: at the beginning of the game ($t=1$), the strategy s is decided and, at that very time the coalition C plans a deviation s_C' to s that will start at period $t>2$ (and is credible after that period) so that it will not become observable until that time. Yet, according to the definition of Com-PE, for this deviation to be consistent, it has to induce a Com-PE at period 2. If some coalition T crediblely deviates from (s_{-C}, s_C') at some time between periods 2 and t, then s_C' was not credible itself and s may be an equilibrium after all; but this way of rulling out a deviation does not seem reasonable. Is this a problem for the definition of Com-PE? The answer is no. This is bacause one has to check any possible deviation from s, and, in our case, there is always a credible one, namely s_C' proposed at period t. Conversely, if s_C' is not ruled out before period t, it cannot be ruled out at period t either. This argument shows that there is no lost of generality if it is assumed that deviations are planned in the same period they are implemented (the definition of Com-PE will be equivalent).

5. TWO EXAMPLES

The purpose of this section is to explore further the heuristics of the definition of Com-PE. The first example deals with the requirement of having an equilibrium after deviations. The second example shows how communication rules out deviations that players rationalize using "forward induction" arguments.

Consider, then, the example in figure 3.

FIGURE 3

Let p_i denote the probability of choosing L_i by player i in a behavioral strategy. Note that, in the subgame after L_1, players 2 and 3 play the "matching pennies" game and that the only equilibrium at this subgame is $(p_2, p_3)=(1/2, 1/2)$. Therefore the Com-PEa are as follows:

(i) if $x<1$, $(p_1, p_2, p_3) = (1, 1/2, 1/2)$;

(ii) if $x>1$, $(0, 1/2, 1/2)$, and

(iii) if $x=1$, $(p_1, 1/2, 1/2)$ with $1 \geq p_1 \geq 0$.

See that, in this example, the set of Com-PEa coincides with that of

SPEa.

Case (i) presents no problem. But, in case (ii) one may argue that, given p_2=1/2, players 1 and 3 may agree to play $(p_1, p_2) = (1, 1)$, an equilibrium in the induced game that gives both players a better payoff (in fact, this is enough to rule out (0, 1/2, 1/2) as a PCPNE). However, player 2 observes the deviation and, in a first approach, one may say that he will try to deduce something about player 3's strategy and play accordingly; he may anticipate p_3=1 and play p_2=1 himself. But once we have open the door to this kind of arguments, it is difficult to stop; player 3 can anticipate player 2's deduction and so on. In other words, following this line of resasoning, one has to accept that, after the deviation by player 1, anything can happen (players 2 and 3 will play rationalizable strategies (Bernheim, 1896)). If communication may take place, there is a better way to analyze the game; according to the scheme in section 4, player 2 just calls for renegotiations to impose the equilibrium (1/2, 1/2) at the subgame. It is now the last agreement for player 3 and the deviation at the beginning was not profitable for player 1.

For the case x=1, the question may not be that simple. Take the (Com-P) equilibrium (1/2, 1/2, 1/2), again players 1 and 3 may deviate to (1, 1). If player 2 observes the randomization by player 1, we can argue as in case (ii); but, if in the more realistic case in which playe 2 only observes the realization of the random choice, he can not conclude anything when he observes that L_1 has been chosen and, then, the deviation $(p_1, p_3) = (1, 1)$ may seem credible[3]. But, even in this case, nothing prevents player 2 from calling for renegotiations just to make sure that (1/2, 1/2)

[3] In this case, the only PCPNE is (1, 1/2, 1/2).

will follow afterwards. At the beginning of the game, players not only agree to play some strategy profile, but they agree to agree to play that strategy at any time they are called for renegotiations (and they will obey the agreement if it is an equilibrium).

For the second example, consider the game of "battle of sexes" in figure 4.

FIGURE 4

If player has the opportunity to "burn a dollar" before playing, it is well known that there is only one stable equilibrium in which player 2 does not burn the dollar and (t,l) follows. With communication, however, (b,r) is still plausible: if (b,r) is decided and player 2 burns the dollar to show that he will play agressively afterwards (to induce (t,l)); at the begining of the second stage he will hear from player 1 something like:

"Ok, you burnt a dollar, so what? we planned to play now (b,r) and we shall do that way since it is an equilibrium; your deviation is worthless so you better follow the equilibrium path."

Still, van Damme argues that, "if the requirement (of Pareto perfectness (= Com-PE in this example)) is really compelling, then ... players should accept the same concept also in the case in which no such communication is possible, especially in the case there is a unique PPE". (van Damme Oct, 1987). Then he shows that the only PPE may be ruled out by fordward induction arguments (stability in the sense of Kolhberg and Mertens (1986)). A more detailed discussion on the relation between stability and renegotiation-proof will be presented in a forthcoming work by the author[4]

6. EXTENDING Com-PE TO INFINITE GAMES

This section will follow the approach by Greenberg (1989) and Asheim (1988) where the von Neumann and Morgenstern abstract stable set is used to extend recursive definitions to the infinite case. The reader is referred to those works for a discussion on this approach.

A von Neumann and Morgenstern abstract system (AS) is a pair $(D,<)$ where D is an abstract set and $<$ is a dominance relation. $d<f$ will be interpreted to mean that f dominates d. Let $(D,<)$ be an abstract system, and let $f \in D$. The dominion of f, denoted by $\Delta(f)$, is the set $\Delta(f)=\{d \in D \ / \ d<f\}$. That is, $\Delta(f)$ consists of all elements of D that f dominates, according to the dominance relation $<$. Similarly, for a subset $F \subset D$, the dominion of F, denoted by $\Delta(F)$, is the set $\Delta(F)=\cup\{\Delta(f) \ / \ f \in F\}$. That is, an element d in D belongs to $\Delta(F)$ if it is dominated by some element in F. A set $F \subset D$ is a von Neumann and Morgenstern abstract stable set (ASS) for the system $(D,<)$ iff $F=D \backslash \Delta(F)$.

Let Γ be a multi-stage game. Inspired by definition 4 and remark 2 an abstract system $(D,<)$ is introduced; let the elements of the abstract set consist of a coalition, a subgame and a SPE in this subgame,

---

[4] Roughly speaking, it will be argued that, in the very particular case when players may belive in the theories of both stable and renegotiation-proof equilibria, then the definition of Pareto perfect equilibrium needs to be changed in a situable way to accomodate the two concepts.

$$D=\{(C,g^h,s) \;/\; C\subset N,\ C\neq\emptyset,\ h\in H,\ s\in E^h\}$$

where E^h is the set of SPE after h; and let the domination relation be defined as follows:

$(C,g^h,s) < (T,g^k,y)$ if and only if either

(i) $k\in H^h$ and $k\neq h$: $T\subset N$, $s_{-T}=y_{-T}$, $y >_T s$ or

(ii) $k=h$: $T\subset C$, $s_{-T}=y_{-T}$, $y >_T s$.

REMARK 5. $(D,<)$ reduces to the AS used in Greenberg's extension of CPNE when $t=1$ and to the AS in Asheim's extension of PPE when $n=2$.

The next proposition relates the ASS of $(D,<)$ with the definition of Com-PE for finite games and allows for a definition applicable to infinite games.

PROPOSITION 5. Let K be an ASS of $(D,<)$; then, for finite games (the number of players, stages and alternatives is finite), we have that for all $h\in H$, $C\subset N$ and $s\in S$;

$(C,g^h,s)\in K$ iff $(s_C(t(h)),s(h_1))$ is a Com-PE in the subgame $g^h|s_{-C}(t(h))$;

in particular, when $C=N$ and $h=\emptyset$, $\text{Com-PE}(\Gamma)=A=\{s/(N,\Gamma,s)\in K\}$.

The proof will follow after two lemmas.

LEMMA 1. For all $h \in H$, $C \subset N$ and $s \in S$; $(s_C(t(h)), s(h_1))$ is a Com-PE in $g^h | s_{-C}(t(h))$ if and only if there are no T, k and x such that $(x_t(t(k)), x(k_1))$ is a CPE in $g^k | x_{-T}(t(k))$ and neither

(i) $k \in H^h$, $T \subset N$, $s_{-T} = x_{-T}$, $x >_T s$; nor

(ii) $k = h$, $T \subset C$, $s_{-T} = x_{-T}$, $x >_T s$

is satisfied.

PROOF. By induction in the number of stages t. If $t=1$, the proof reduces to show that for all $C \subset N$, s_C is a Com-PE (=CPNE) in $\Gamma | s_{-C}$ implies that no $T \subset N$ and $x \in S$ exist such that x_T is CPNE in $\Gamma | x_{-T}$ and $T \subset C$, $x_{-T} = s_{-T}$, $x >_T s$; but this comes from lemma 1 in Greenberg (1989).

Assume that the lemma has been proved for games of less than t stages and prove it now for $t(>1)$. If $(s_C(t(h)), s(h_1))$ is a Com-PE in $g^h | s_{-C}(t(h))$ and $k \in H^h$ and if $k \neq h$ then $s|_k$ is Com-PE in g^k by definition of Com-PE; apply the induction hypothesis to get that no $T \subset N$ and $x \in S$ exist such that $(x_T(t(k)), x(k_1))$ is a Com-PE in $g^h | x_{-T}(t(k))$ with $x_{-T} = s_{-T}$ and $x >_T s$. Finally, for the case $k=h$, supose that there exist $T \subset N$ and $x \in S$ such that $(x_T(t(h)), x(h_1))$ is a Com-PE in $g^h | x_{-T}(t(h))$, then x_1 is a Com-PE in $g^{h(1)}$; if $T \subset C$, $s_{-T} = x_{-T}$ and $x >_T s$ then s was not a CPNE restricted to $S' = \{s \in S$ / s induces a Com-PE in proper subgames of $\Gamma\}$ and the proof is complete.

LEMMA 2. For all $h \in H$, $C \subset N$ and $s \in S$, if $(s_C(t(h)), s(h_1))$ is <u>not</u> a Com-PE then <u>there exist</u> $T \subset N$, $k \in H$ and $x \in S$ such that $(x_T(t(k)), x(k_1))$ is a Com-PE in $g^h | s_{-T}(t(k))$ and either

(i) $k \in H^h$, $k \neq h$, $s_{-T} = x_{-T}$ and $x >_T s$ or

(ii) $k=h$, $T\subset C$, $s_{-T}=x_{-T}$ and $x >_T s$.

PROOF. If $t=1$, it comes from lemma 1 in Greenberg (1989).
If $t>1$, that $(s_C(t(h)), s(h_1))$ is not a Com-PE in $g^h|s_{-C}(t(h))$ means that either

(i') s_1 is not a Com-PE in g^k with $k\neq h$, $k\in H^h$, or

(ii') s_1 is a Com-PE in $g^{h(1)}$ but there exist $T\subset C$ and $x_T\in S_T$ such that $(x_T, s_{-T}) >_T s$ and (x_T, s_{-T}) is a Com-PE in $g^h|_{-T}(t(h))$.

(i') implies (i) and (ii') implies (ii).

PROOF OF PROPOSITION 5. Comes from lemmas 1 and 2.

Since the characterization of a Com-PE given in proposition 5 is not based on backwards recursion, it can be used to formulate a general definition of the concept, covering both finite and infinite horizon games and games with either a finite or an infinite number of players.

DEFINITION 7. Consider a multi-stage game Γ with a finite or infinite number of players. A strategy profile s is said to be a Com-PE of Γ if and only if there is a ASS, F, for the associated system $(D,<)$ such that $(N,\Gamma,s) \in F$.

F is interpreted as a "social norm"; every point in F is equally reasonable: no one dominates any other in the social norm and any point outside the social norm is dominated by some element inside it.

APPENDIX

PROOF OF PROPOSITION 1. For $t=1$ SPE = NE = SPE′ trivially. Assume that the proposition has been proved for games of less than t stages, and let Γ be a game of $t>1$ stages. Let $s^*\in SPE$ and let $T = \{s\in S \,/\, s$ induces a SPE in propoer subgames of $\Gamma\} = \{s\in S \,/\, s$ induces a SPE′ in proper subgames of $\Gamma\}$, by induction hypothesis. Clearly $s^*\in T$ and s^* is a NE, which implies that s^* is a NE restricted to T and by definition 3′ this means that s^* is a SPE′. Now let s^* be a SPE′, this means that s^* is a NE restricted to T as defined above. Then s^* is a SPE in all subgames with less than t stages. If s^* is not a NE there exists a strategy $s_i\in S_i$ such that $u^i(s^*_{-i}, s_i) > u^i(s^*)$. Consider two cases:

(1) $s_i(g)=s^*_i(g)$ where g is a proper subgame of Γ. Then s^* is not a NE restricted to T $(s^*\notin T)$, and the proposition is proved.

(2) $s_i(g)\neq s^*_i(g)$ for some proper subgame g. The induced outcome of g under (s^*_{-i}, s_i) is defined by $u(s^*_{-i}(g), s_i(g))$; then it has to be that $u^i(s^*_{-i}(g), s_i(g)) \leq u^i(s^*(g))$ since s^* is a NE in every subgame. Next consider

$$s''_i = \begin{cases} s'_i & \text{in the first stage of } \Gamma \\ s^*_i & \text{in proper subgames of } \Gamma \end{cases}$$

clearly $u^i(s^*_{-i}, s''_i) > u^i(s^*)$; then, by case (1) s^* is not a NE restricted to T.

PROOF OF PROPOSITION 5 (iv). Consider the game in figure 5

FIGURE 5

There are two SPEa in the game:

(i) $(l_1, l_2, l_3, l^*_3, l^*_1, l^{**}_1, l^*_2, l^{**}_2)$ and

(ii) $(r_1, r_2, r_3, l^*_3, r^*_1, l^{**}_1, r^*_2, l^{**}_2)$.

The equilibrium in (i) is Com-PE but not PCPNE. To show that it is Com-PE see that no coalition of n players can deviate to a better stragtegy (n=1,2,3). If n=1, because (i) is SPE; if n=3, because the payoffs induced in every subgame are Pareto optimal; if n=2, see that only the coalition formed by players 2 and 3 can find an outcome in which both players are better off, namely the outcome (0,9,5) which is preferred to (8,4,4). But in order to obtain that outcome, the deviation to (r_2, r^*_3, r^{**}_2) has to take place in the second stage of the game, but that deviation does not conform a Com-PE in the subgame after r_2. Since that deviation is a PCPNE (it is SPE and optimal) in the game induced by fixing player 3's strategy (l_1, l^*_1, l^{**}_1), we conclude that (i) is not a PCPNE.

Equilibrium (ii) is PCPNE but not Com-PE. It is not Com-PE because it is Pareto dominated by (i). It is a PCPNE because it induces a PCPNE in every proper subgame (it is straightforward to check), because it induces a PCPNE in games induced by player i's strategies (also straightforward since only subgame perfection and optimality is needed) and because it is not dominated by any other PCPNE.

REFERENCES

Abreu, D. and D. Pearce (1988): "A Perspective on Renegotiation in Repeated Games", Working Paper in Economics E-89-31, The Hoover Institution, Stanford University.

Asheim, G. B. (1988): "Extending Renegotiation-Proofness to Infinite Horizon Games", mimeo, The Norwegian School of Economics and Business Administration.

Aumann, R. (1959): Acceptable points in general cooperative n-person games, in "Contributions to Theory of Games IV," Princeton Univ. Press, Princeton, N.J., 1959.

Bernheim, B. D. (1986): "Rationalizable Strategic Behavior", Econometrica, 1007-1028.

Bernheim, B. D. and D. Ray (1989): "Collective Dynamic Consistency in Repeated Games", Games and Economic Behavior, 1, 295-326.

Bernheim, B. D; Peleg B. and M. D. Whinston (1987): "Coalition-Proof Nash Eqilibria I. Concepts". Journal of Economic Theory, 42, 1-12.

Farrell, J. and E. Maskin (1989): "Renegotiation in Repeated Games", Games and Economic Behavior, 1, 327-360.

Greenberg, J. (1989): "Deriving Strong and Nash Equilibria from an Abstract System", Journal of Economic Theory, 49, 195-202.

Kolhberg, E. and J. L. Mertens (1986): "On the Strategic Stability of Equilibria"; Econometrica 54, 1003-1037.

Peleg, B. (1988): "On Perfectly Coalition-Proof Nash Equilibria", mimeo, Department of Mathematics, The Hebrew University.

van Damme, E. (October 1987): "The Impossibility of Stable Renegotiation"; Discussion Paper No. A-136, SFB 303 Universitat Bonn.

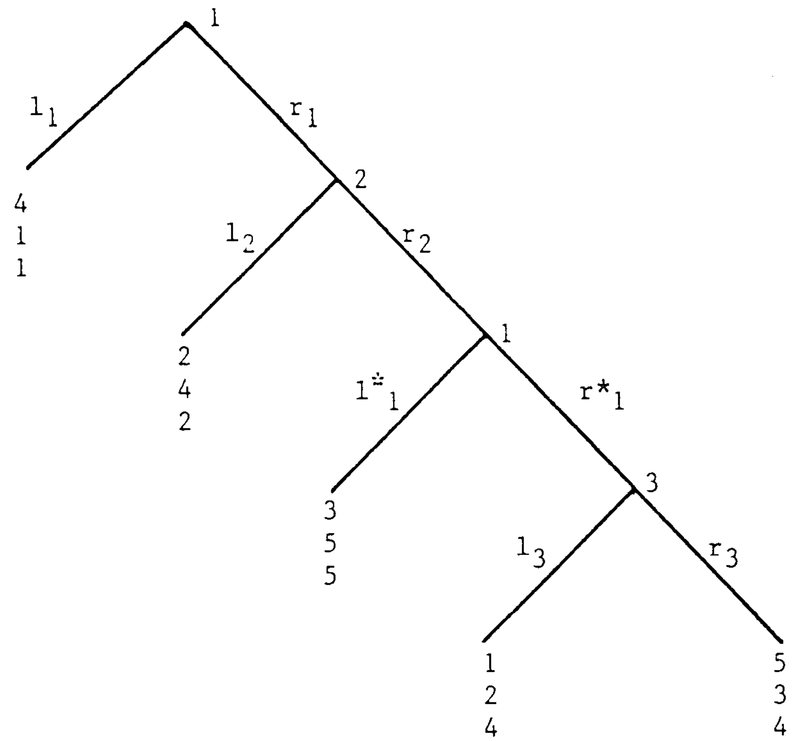

Figure 1

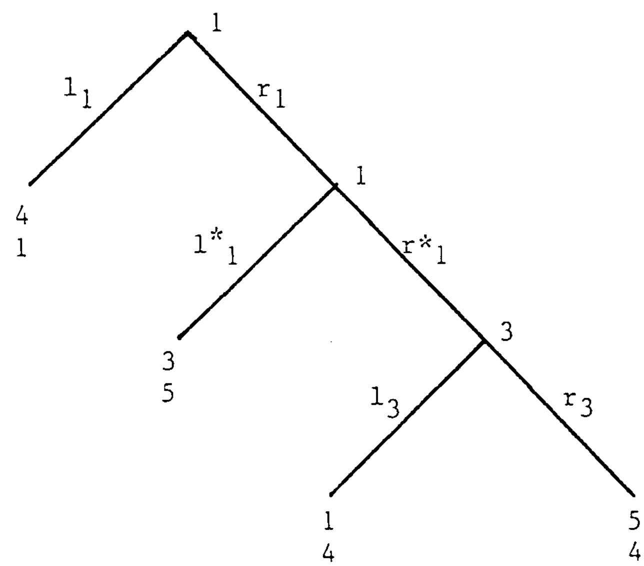

Figure 2

Player 2

		l	r
Player 1	t	2, 2	0, 0
	b	0, 0	3, 3

Figure 3

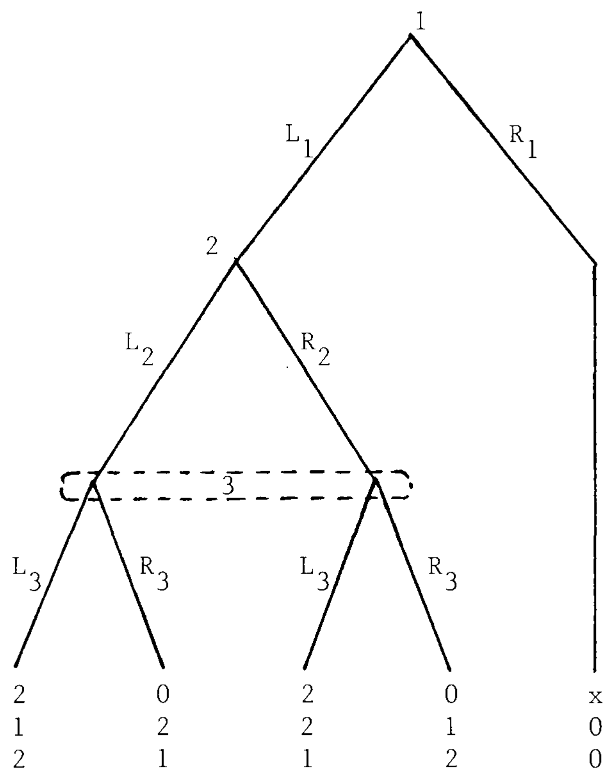

Figure 4

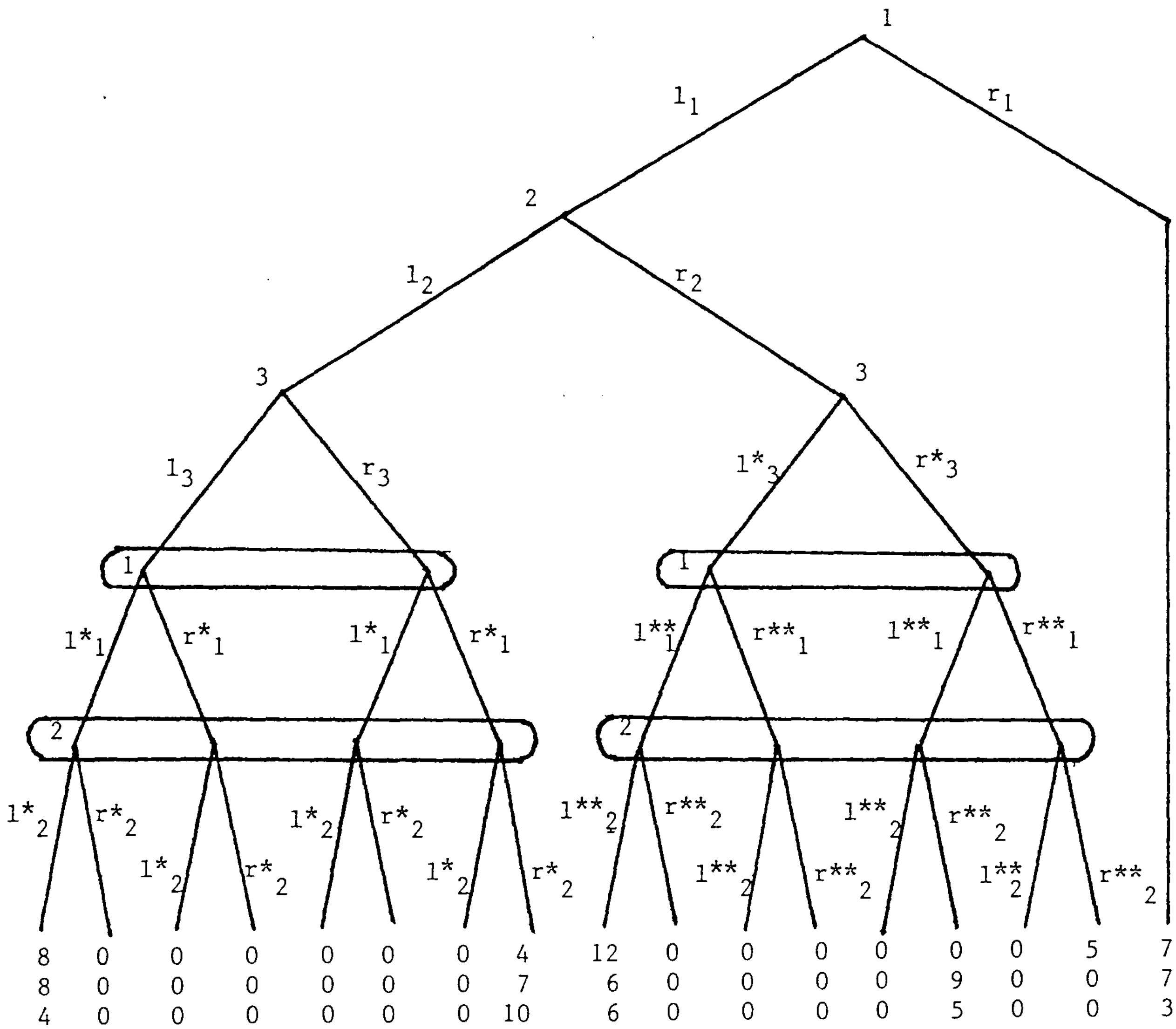
1
l_1
r_1
2
l_2
r_2
3
3
l_3
r_3
l^*_3
r^*_3
1
1
l^*_1 r^*_1 l^*_1 r^*_1
l^{**}_1 r^{**}_1 l^{**}_1 r^{**}_1
2
2
l^*_2 r^*_2 l^*_2 r^*_2 l^*_2 r^*_2 l^*_2 r^*_2
l^{**}_2 r^{**}_2 l^{**}_2 r^{**}_2 l^{**}_2 r^{**}_2 l^{**}_2 r^{**}_2
8 8 4
0 0 0
0 0 0
0 0 0
0 0 0
0 0 0
0 0 0
4 7 10
12 6 6
0 0 0
0 0 0
0 0 0
0 0 0
0 9 5
0 0 0
5 0 0
7 7 3

Figure 5