

A PERSPECTIVE ON RENEGOTIATION IN REPEATED GAMES

by

Dilip Abreu and David Pearce*

Abstract: In this paper we discuss the conceptual foundations of one approach to modelling renegotiation in repeated games. Renegotiation-proof equilibria are viewed as social conventions that players continue to find beneficial after every history. The theory can be understood in terms of stationary stable sets of credible deviations.

Self-enforcing agreements play an important role in many areas of economics, and in the social sciences more generally. An extensive literature uses the repeated game model to explore formally what can be achieved by self-enforcing agreements among rational individuals. In a serious challenge to this work, a number of authors beginning with Farrell (1984) have argued that most of the agreements previously considered are vulnerable to *renegotiation*: after some histories of play, the participants will scrap the original implicit contract, and forge a new agreement. While the validity of the attack is widely accepted, the appropriate definition of a renegotiation-proof equilibrium is less clear. The most widely explored approach has been pursued by Farrell and Maskin (1987), Bernheim and Ray (1987), van Damme (1986), Blume (1988), Bernheim and Whinston (1986), and Benoit and Krishna (1988). Other possibilities are studied by Cave (1986), Asheim (1988), DeMarzo (1988) and Bergin and MacLeod (1989). Most of these papers focus on definitions and characterizations, with less space devoted to discussing the conceptual foundations for the respective definitions. Here we try to explain the thinking behind the approach we have taken in our own work on renegotiation (Pearce (1987) and Abreu, Pearce and

*We would like to thank Reinhard Selten for encouraging us to write this essay and our co-author Ennio Stacchetti with whom some of the ideas presented here were developed. We are grateful for the hospitality of the ZiF at the University of Bielefeld, and for the financial assistance of the Sloan Foundation and the National Science Foundation.

Stacchetti (1989)) and to pinpoint the differences in perspective that distinguish it from the standard formulations in the literature.

Logically prior to the issue of renegotiation—proofness is the question of whether *any* cooperation among players is consistent with their rationality. We touch briefly on this subject before describing the Farrell—Maskin (1987) definition of a weakly renegotiation—proof set. A discussion of the contrasting reasoning that shapes their formulation and ours, respectively, leads us to suggest a framework that unifies several of the concepts introduced in Pearce (1987) and Abreu, Pearce and Stacchetti (1989). This interpretation of our work shows the influence of Greenberg (1988) and Asheim (1988). Finally, we apply our line of reasoning to finitely repeated games. Throughout, discounted repeated games with perfect monitoring are the subject of inquiry, but the same considerations apply to simple imperfect monitoring models.

COOPERATION AND SELF—ENFORCING AGREEMENTS

Suppose that some simultaneous game G is repeated indefinitely, and denote by $G^\infty(\delta)$ the resulting supergame in which each player's payoff is the discounted sum (according to the discount factor δ) of the infinite stream of his payoffs in the component games. An equilibrium of $G^\infty(\delta)$ is said to display *cooperation* if, at some point on the equilibrium path, behavior does not correspond to any Nash equilibrium of the stage game G . In other words, in some period at least one player is not playing a myopic best response to others' choices, presumably because he expects other players' future choices to vary according to his current play. Is such variation consistent with rationality?

Güth, Leininger and Stephan (1988) argue for the imposition of subgame consistency (see Selten (1973) or Harsanyi and Selten (1988)) on solutions of infinitely repeated games. This principle requires that if two subgames are isomorphic, behavior of rational players should be identical in the two games. Notice that all subgames of $G^\infty(\delta)$ are isomorphic:

each is a perfect copy of $G^\infty(\delta)$ itself. Thus, subgame consistency rules out the variation in behavior that might permit cooperation to arise.

Rationality in a strategic context is a notoriously delicate concept, and the term has been used in many ways. Certainly the theorist can impose as an axiom the doctrine that the internal structure of any given game (or subgame) determines a unique profile that is consistent with player rationality; from this starting point, one can build clean and tractable theories. But we do not see the necessity for such an approach, nor does it seem the most attractive way to idealize intelligent human behavior. Consider a social convention, or arrangement, that creates the presumption that certain actions will be followed by particular responses. If this gives incentives that allow players to cooperate to their mutual profit, greedy players might be expected to adopt the convention and perpetuate it. Of course, not all arrangements can inspire confidence; an agreement should be self-enforcing. Exactly what this means is admittedly a difficult question. Subgame perfection (Selten (1965, 1975)) is a necessary feature of a self-enforcing agreement, but it does not consider the possibility that the group (as opposed to an individual) might in concert abandon the equilibrium plan. This is precisely the issue that the literature on renegotiation in supergames addresses.

RENEGOTIATION AND CREDIBLE DEVIATIONS

A theory of renegotiation attempts to identify those equilibria that can survive the threat of renegotiation in every contingency. The definition of a weakly renegotiation-proof set given by Farrell and Maskin (1987) is representative of the most popular approach in the literature. A *weakly renegotiation-proof set* is "... a collection S of perfect equilibria and all their continuation equilibria, such that at no point in the game tree of any equilibrium in S would players all prefer moving to another member of S ." (Farrell and Maskin (1987), page 3) The elements of S are interpreted as the agreements the players might credibly make.

Think for a moment of the particularly simple case in which G is a symmetric game and all of the subgame perfect equilibria of $G^\infty(\delta)$ are strongly symmetric (all players expect the same payoffs as one another in any given subgame), or in which attention is restricted to such equilibria. Here a weakly renegotiation-proof set cannot contain two equilibria with different values, because they would be Pareto-ranked. Since continuation payoffs are constant, there can be no cooperation in this setting.

This seems to us unnecessarily pessimistic. Suppose G has a unique Nash equilibrium with value 0, and $G^\infty(\delta)$ has subgame perfect equilibria with average discounted values ranging from -4 to 5 , for example. Suppose further that there is an equilibrium σ all of whose continuation values lie between 2 and 4 ,¹ inclusive, and that *all* perfect equilibria of $G^\infty(\delta)$ include continuation values of 2 or less in *some* contingencies. We would consider σ to be renegotiation-proof, for the following reason. If, in a subgame in which the continuation value of σ is 2 , someone proposes renegotiation to an equilibrium γ with a higher value, he is implicitly making the case that the value 2 is needlessly severe, and can be abandoned. If it can be abandoned now, *in violation of a prior agreement*, it can presumably be abandoned in any future contingency. But γ (like all equilibria of $G^\infty(\delta)$) specifies the value 2 or worse in some subgame, so γ itself is vulnerable to renegotiation. In other words, the suggestion that the group can credibly renegotiate away from 2 is internally inconsistent. To say that an equilibrium with value 4 may be credible as part of an ongoing plan, is *not* to imply that it is believable as a proposed *deviation*.

Notice that Farrell and Maskin (1987), Bernheim and Ray (1987) et al. do not distinguish between the credibility of an ongoing plan, and its credibility as a *breach* of a social convention. We favor making the distinction, both on logical grounds and for its intuitive appeal. It is easy to give examples drawn from common experience illustrating

¹As Pearce (1987) and Abreu, Pearce and Stacchetti (1989) show, equilibria all of whose continuation values lie strictly above the Nash value, are not exceptional.

the operation of the distinction. Society threatens jail sentences for embezzlement, and routinely carries out the threat (which is costly to all concerned). Not jailing the innocent is also part of the convention. No one concludes that since it is credible (in a society that needs to deter embezzlement) to let a person go free if he has committed no crime, it must also be credible not to jail embezzlers. The latter act of setting free is a breach of the social code, and undermines the public's belief in the code's relevance. Similarly, it is credible for a parent both to reward a child's good behavior and to punish bad behavior. Punishment is necessary since the child cannot take seriously the idea that although misbehavior has gone unpunished in the past, in the future punishment will be sure and swift.

The principal objection to the distinction we propose is that it requires a lack of stationarity in the set of strategy profiles that are credible: a given profile is believable after some histories, but not after others. Most of the literature on renegotiation assumes that because the repeated game is stationary, so too is the set of things that are credible or plausible. Carried to its logical conclusion, *this argument rules out cooperation entirely*. Suppose that anything that is plausible or credible at one point in the supergame is also credible in any other contingency. Presumably this applies to strategy profiles, suggestions, threats, promises and so on. Since each player has access to the same negotiation statements after some history h as he had at the beginning of the game, and exactly the same ones are plausible, each player should be able to secure the same continuation payoff as he negotiated at the beginning of $G^\infty(\delta)$. But if no one's continuation payoff ever varies, there are no incentives for cooperation.

We suggest that the natural way to impose stationarity is to require that the set of credible *deviations* is always the same. A credible deviation is a strategy profile to which the group believes it could defect, in *violation* of the previously negotiated plan. Recall that a history of play has no "physical" relevance in the subgame that follows it; it matters

only via the social convention. Since a deviation by the group finds no support in the convention it is *breaching*, the credibility of the deviation can hardly be history-dependent.

Even if the group believes it can deviate to a particular profile, it may not want to do so. As usual, it is hard to determine what a group "wants." Various criteria may be used to decide whether or not the deviation will occur; the Pareto criterion is the one usually invoked in the literature. We shall return to a discussion of this and alternative criteria later. For the moment suppose that \succ is a partial ordering of the set of supergame profiles, with the following interpretation. If x is a profile specified by the social convention in force, and y is a credible deviation with $y \succ x$ (read " y dominates x "), then the group will abandon x in favor of y . Thus, a self-enforcing agreement must have the property that none of its continuation equilibria is dominated by any credible deviation. It is also desirable to explain why a profile is *not* considered a credible deviation after some history it is dominated by some credible deviation. The last two remarks suggest a formalization in terms of *stable sets* of credible deviations, in the spirit of von Neumann and Morgenstern (1947). Greenberg (1988) has emphasized the applicability of the stable set approach to diverse strategic problems. Asheim (1988) pursues Greenberg's lead in the context of repeated games and renegotiation-proof equilibria. The novelty in our work is the distinction we make between ongoing agreements and credible *deviations*. This leads to quite a different theory from Asheim's, the latter being more in the spirit of Farrell and Maskin (1987), Bernheim and Ray (1987) et al

STABLE SETS OF CREDIBLE DEVIATIONS

Let S be the set of pure strategy profiles of G , and Σ be the set of pure strategy profiles of $G^\infty(\delta)$. For any $\sigma \in \Sigma$ and history h , $\sigma|h$ denotes the profile induced by σ on the subgame following h . Fix a (strict) partial ordering Ω of Σ . For each set of supergame strategy profiles $\Omega \subseteq \Sigma$, define

$$D(\Omega) = \{\sigma \in \Sigma \mid \exists h \in \cup_t S^t \text{ and } \gamma \in \Omega \text{ such that } \gamma \succ \sigma|h\},$$

the set of profiles dominated after some history by elements of Ω . For singleton sets $\{\sigma\}$ we abuse notation by writing $D(\sigma)$.

An equilibrium strategy profile σ cannot be credible as a deviation if it dominates one of its own continuation equilibria: the continuation equilibrium would be abandoned in favor of σ . Consequently, we confine attention to the set of *consistent deviations*:

$$\Gamma = \{\sigma \in \Sigma \mid \sigma \text{ is a subgame perfect equilibrium and } \sigma \notin D(\sigma)\}$$

We are now ready to define a set of credible deviations, and a renegotiation-proof equilibrium associated with the set.

DEFINITION: A set Ω of consistent deviations is a *set of credible deviations* if

- (i) $\Omega \cap D(\Omega) = \emptyset$ Internal Stability
- (ii) $\Gamma \setminus \Omega \subseteq D(\Omega)$ External Stability

Nothing in Ω is dominated by any element of Ω , whereas all consistent deviations not in Ω are dominated from within Ω .

DEFINITION: The set of equilibria that are *renegotiation-proof* relative to a credible set of deviations Ω is $R(\Omega) = \{\sigma \in \Sigma \mid \sigma \text{ is a subgame perfect equilibrium and } \sigma \notin D(\Omega)\}$.

Notice that the difference between Ω and $R(\Omega)$ is that the elements of $R(\Omega)$ are not required to be consistent. Recall our distinction between credibility as part of an ongoing plan, and credibility as a deviation. An element of $R(\Omega)$ may dominate one of its continuation equilibria, say $\sigma|h$. This does not threaten the credibility of σ , because the

group does not believe it can *deviate* to σ from an ongoing agreement. If it believed it could, it would also deviate from $\sigma|h$ to σ ; this shows the internal inconsistency of believing σ to be a credible deviation.

The specific theory that emerges from this framework depends on the choice of \succ , the dominance relation. In the special environment in which G is symmetric and attention is restricted to strongly symmetric equilibria of $G^\infty(\delta)$, \succ should simply order the equilibria according to the players' unanimous ranking. The resulting theory coincides with the definition given by Pearce (1987): an equilibrium σ is renegotiation-proof if the infimum of its set of continuation values is at least as high as the corresponding infimum for any other equilibrium. As long as the threat of permanent reversion to myopic Cournot-Nash behavior would suffice to sustain some cooperation, this theory (unlike the standard ones) also predicts cooperation (Pearce (1987)). Thus, even in this setting where there is no argument about which dominance criterion to use, solutions depend critically on whether or not deviations are distinguished from ongoing plans.

When players are not unanimous about the ranking of a continuation equilibrium and a credible alternative, most authors assume that the deviation will occur if and only if all players strictly prefer it to the status quo. Letting the relation \succ in the framework above be the Pareto partial ordering is not compatible with the existence of solutions. Pearce (1987) proves the existence of a weaker solution that combines the Pareto rule with the idea of credible deviations, but does not impose external stability. Similarly, Farrell and Maskin (1987) et al. use the Pareto partial ordering in the standard formulation without requiring external stability. Asheim (1989) imposes both internal and external stability with a Pareto dominance criterion, but does not prove existence, showing instead that stationarity is incompatible with existence.

Despite its popularity, the Pareto rule seems a dubious criterion for predicting the occurrence of renegotiation. In this respect it is a convincing sufficient condition: if everyone strictly prefers a credible alternative to the status quo, the latter is sure to be

abandoned. But to require unanimous consent before a deviation occurs is extreme. In particular, the Pareto criterion is in conflict with one's sense that in situations with multiple equilibria, the outcome should depend on the bargaining positions of the participants. In general this involves highly complex considerations.

Abreu, Pearce and Stacchetti (1989) (hereafter APS) introduce a bargaining approach while sidestepping some of its difficulties by focusing on symmetric games. Although a multitude of factors may contribute to bargaining power, presumably this is the same for each player. In a static setting, equal bargaining power might be expected to lead, under modest assumptions, to a symmetric outcome. Similarly, in the repeated game, a player might reasonably insist on "equal treatment," except to the extent that allowing relative payoffs to vary with the history is in his interest. An arrangement that exploits variations in shares is credible only if a player finds the convention useful even when his fortunes are lowest. We assume that an individual will not object to a continuation payoff of w , say, if in *every* subgame perfect equilibrium, someone must accept w or less in some subgame. This motivates the solution concept studied by APS. A subgame perfect equilibrium σ is a *consistent bargaining equilibrium* if the lowest continuation payoff received by any player in any subgame is at least as high under σ as for any other subgame perfect equilibrium. Intuitively, this means that after no history can anyone protest that he is suffering an unnecessarily severe punishment.

In the framework introduced earlier, let \succ be the Rawlsian ordering: for any σ, γ, ξ $\sigma \succ \gamma$ if and only if $\min_i v_i(\sigma) > \min_i v_i(\gamma)$, where for any $\xi \in \Sigma$, $v_i(\xi)$ is player i 's associated (average) repeated game payoff. One can check that there is a unique set Ω of credible deviations and that $R(\Omega)$ is exactly the set of consistent bargaining equilibria. Existence is guaranteed by a proposition we establish below. In games with imperfect monitoring, the flexibility gained by exploiting asymmetric continuation rewards can be dramatic, and hence a consistent bargaining equilibrium can easily be asymmetric. But APS show that in games of perfect monitoring satisfying some regularity assumptions, solutions are always

strongly symmetric. This result leads to entirely elementary characterizations of maximal credible collusion in terms of the data of the stage game G .

The drawbacks of using the Rawlsian ordering as a dominance criterion are that it is a somewhat extreme way to model bargaining power, and that it is quite unconvincing in naive extensions to asymmetric games. An alternative that comes immediately to mind is the Nash bargaining solution (Nash, 1950). In cases in which some equilibrium of the stage game provides a natural threat point, letting \succ be the ordering induced by the Nash product yields an interesting theory. Since the priorities of the group are summarized by the Nash solution, renegotiation will occur in any contingency in which the Nash product is "unnecessarily" low. More precisely, for each equilibrium of $G^\omega(\delta)$, one can compute the Nash product of the differences between players' average discounted payoffs and their respective threat point payoffs. A subgame perfect equilibrium σ withstands potential renegotiation if the minimum value of the Nash product, over all subgames, is at best as high as the corresponding minimum for any other perfect equilibrium \succ . Solutions in this sense always exist, as the theorem given below establishes.

A closer look reveals that, despite its attractive features, this theory is less than satisfactory. It is sabotaged by the inadequacy of the Nash bargaining solution as an expression of players' bargaining power in this context. To see this vividly, think of a "dummy player" whose actions have *no* effects on others' payoffs, but whose payoffs are affected by others' choices. Two games differing only by the addition of a dummy player can have quite different Nash bargaining solutions, despite the fact that the dummy player has no bargaining power. This is unacceptable in a positive theory of self-interested strategic behavior.

To our knowledge, no solution concept has yet been formulated that appropriately summarizes players' relative bargaining positions in a repeated game. One could at best hope for a concept that would rank equilibria on the basis of attractive principles; unanimity amongst theorists regarding the most appealing ranking is unlikely to be

achieved. Thus, we are not able to propose a particular theory of renegotiation in general repeated games. Rather, we are advocating an *approach* to the subject, one that is formalized in the "stable sets of credible deviations" framework presented earlier. If in some context there is a dominance rule \succ that seems suitable, the framework yields a corresponding set of renegotiation-proof equilibria of the infinite horizon game.

We prove the existence and uniqueness of a (stable) set of credible deviations when the strict partial ordering \succ is derived from a continuous,² complete ordering R on \mathbf{R}^n . (That is, $\sigma \succ \gamma$ if and only if $[v(\sigma) R v(\gamma)$ and $\sim(v(\gamma) R v(\sigma))]$ where $v(\sigma) \in \mathbf{R}^n$ is the repeated game payoff-vector associated with the strategy profile σ). This assumption seems to us to provide the most coherent setting within which to discuss the issue of renegotiation.

The strict partial ordering \succ on Σ is derived from a continuous, complete ordering R on \mathbf{R}^n .

The proof also uses the following standard assumption.

(A2) The strategy sets S_i and the payoff functions $\Pi_i: S_1 \times \dots \times S_n \rightarrow \mathbf{R}$ of the stage game G are compact and continuous, respectively. Furthermore G has a Nash equilibrium in pure strategies.

We confine attention to pure strategy equilibria of $G^\infty(\delta)$. The last part of (A2) serves only to guarantee that the repeated game has a pure strategy equilibrium.

PROPOSITION: Under (A1) and (A2) a unique set of credible deviations exists.

²That is, for all $w \in \mathbf{R}^n$ the sets $\{x \in \mathbf{R}^n | x R w\}$ and $\{x \in \mathbf{R}^n | w R x\}$ are closed.

PROOF (Sketch): The proof uses the idea of self-generating sets and associated results from Abreu, Pearce and Stacchetti (1986), and mimics Proposition 1 of Pearce (1987). The reader who wishes to follow the argument closely will need to consult the latter.

First note that under (A2), the ordering R can be represented by a continuous function $u: R^n \rightarrow R$ (see Debreu (1954)). Let Θ be the set of all compact self-generating sets. For $X \in \Theta$ let $\underline{X} = \min \{u(x) | x \in X\}$. Let $M = \{\underline{X} | X \in \Theta\}$. As in Proposition 1 of Pearce (1987) it follows that there exists $\underline{W} \in M$ such that $\underline{W} = \sup M$. Let Ω be the set of all equilibria σ such that $u(v(\sigma)) = \underline{W}$ and $u(v(\sigma|h)) \geq \underline{W}$ for all $h \in \cup_t S^t$. Let γ be a consistent deviation. Then $u(v(\gamma)) \leq u(v(\gamma|h))$ for all h . Since for any equilibrium γ , the set $C(\gamma)$ of continuation values of γ , is a self-generating set, it follows that $u(v(\gamma)) \leq \underline{W}$. Furthermore $\gamma \in \Omega$ if and only if $u(v(\gamma)) = \underline{W}$. Clearly Ω satisfies internal and external stability and is the unique set of consistent deviations to do so.

Q.E.D.

FINITELY REPEATED GAMES

For any positive integer T , $G^T(\delta)$ denotes a discounted supergame consisting of T successive periods of G . Repeated games of this kind are generally considered the most promising setting for an uncontroversial treatment of renegotiation-proofness. The applicability of backward induction arguments has resulted in virtual unanimity among those writing on renegotiation in finite horizon games. We give the recursive definition by Bernheim and Ray (1987) of a Pareto-perfect equilibrium; the formulations of Benoit and Krishna (1988), van Damme (1986), Farrell and Maskin (1987) and Asheim (1988) are identical or equivalent. An equilibrium of $G^1(\delta)$ is Pareto-perfect if it is not Pareto-dominated by any other equilibrium of $G^1(\delta)$. An equilibrium of $G^T(\delta)$, $T = 2, 3, \dots$, is Pareto-perfect if it is Pareto efficient within the set of subgame-perfect equilibria all of whose continuation equilibria after the first period are Pareto-perfect in $G^{T-1}(\delta)$.

Benoit and Krishna have derived elegant characterizations of the Pareto-perfect set for large T and $\delta = 1$.

We find the consensus unconvincing. Recall that a history-dependent equilibrium is an arrangement unsupported by any structural consequences of play: all subgames of length t are isomorphic. The arrangement is adopted and retained only as a device for sustaining cooperation. But in the final period of a finite horizon game, it is common knowledge among the players that the convention no longer serves this purpose: there are no succeeding periods, so cooperation is impossible. Why, then, should players be guided in the last period by a convention that has outlived its usefulness?

Pareto perfection requires unanimity for renegotiation to occur. Therefore a solution may exhibit behavior in certain final-period subgames of $G^T(\delta)$ that differs dramatically from what one would predict as the probable outcome of the component game G . An example illustrates the point:

		2				
		b_1	b_2	b_3	b_4	
	a_1	8,8	0,9	0,0	0,0	
	a_2	9,0	5,5	6,0	0,-10	
1	a_3	0,0	10,0	7,1	0,0	G
	a_4	0,0	0,6	0,0	1,7	

In the game G shown above, there are three pure strategy Nash equilibria, having respective values (5,5), (7,1), and (1,7). We submit that the equilibrium (a_2, b_2) having value (5,5) is far more likely to be played than is either of the others. This is attractive on the basis of symmetry, and considerations of utilitarianism, riskiness and self-signalling claims (Farrell (1988)) only strengthen the case. By playing a_3 , player 1 risks receiving –

10. Moreover, claiming to play a_3 and exhorting player 2 to respond with b_3 is unconvincing: if player 1 were intending to choose a_2 , he would still want 2 to choose b_3 . Suggesting the profile (a_2, b_2) , on the other hand, is unambiguous.

The reader may check that the following instructions form a Pareto-perfect equilibrium of $G^2(1)$: play (a_1, b_1) in the first period, followed by (a_2, b_2) if no one has deviated or if both have, (a_3, b_3) if player 2 alone deviates, or (a_4, b_4) if player 1 alone deviates. Ex ante, this scheme is attractive, promising high payoffs for both participants. But how credible is the promise? If player 1 were to cheat in the first period, he should subsequently announce his intention to play a_2 in the terminal period. The arguments at his disposal seem likely to prevail over player 2's plaintive reminders about the initial agreement.

Not all games are as easy to analyze as the component game G just discussed. Theorists may find it difficult to predict with any confidence what will occur even in a single static game. But this does not justify the construction of equilibria in which final-period payoffs are manipulated to support cooperation earlier in the repeated game. Behavior in the last period is more likely to be determined by the players' view of their bargaining position in the one-shot game. Immutability of final period play then spreads to the other periods by the usual backward induction arguments.

CONCLUSION

To achieve cooperation in a repeated game, players must adopt a social convention specifying how behavior will vary according to the history of play. Thus, although all subgames of an infinitely repeated game are identical, players' views of what is credible are history-dependent. In particular, a strategy profile may be credible as part of an ongoing arrangement, but not if instead it constitutes a breach of the social convention. Since a deviation is "unauthorized," its credibility is arguably independent of the social convention, and hence of the history. For infinitely repeated games, this suggests that one

look for a stationary stable set of credible deviations; these, in conjunction with a rule summarizing how the group will choose between an ongoing equilibrium and a credible alternative, determine which agreements are actually self-enforcing. We find the Pareto rule unsuitable for this role; what is needed is a criterion reflecting players' bargaining power. The notion of a consistent bargaining solution (APS (1989)) in symmetric games is in this spirit; it can be given a stable set interpretation using the framework developed here. A satisfying definition of renegotiation-proof equilibrium for asymmetric repeated games awaits the development of more sophisticated theories of bargaining for these games.

REFERENCES

Abreu, D., D. Pearce, and E. Stacchetti (1986): "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39, 251–269.

Abreu, D., D. Pearce, and E. Stacchetti (1989): "Renegotiation and Symmetry in Repeated Games," Cowles Foundation Discussion Paper.

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

Benoit, J.–P., and V. Krishna (1988) "Renegotiation in Finitely Repeated Games," mimeo, Harvard Business School.

Bergin, J., and W. B. MacLeod (1989): "Efficiency and Renegotiation in Repeated Games," mimeo, Queen's University.

Bernheim, B.D., and D. Ray (1987): "Collective Dynamic Consistency in Repeated Games," mimeo, Stanford University.

Bernheim, B.D., and M. Whinston (1987): "Coalition–Proof Nash Equilibria II Applications," *Journal of Economic Theory*, 42, pp. 13–29.

Blume, A. (1987): "Renegotiation–Proof Theories in Finite and Infinite Games," mimeo, University of California, San Diego.

Cave, J. (1987): "Long Term Competition in a Dynamic Game: The Cold Fish War," *Rand Journal of Economics*, 18, pp. 596–610.

van Damme, E. (1989): "Renegotiation–Proof Equilibria in Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 47, pp. 206–207.

Debreu, G. (1954): "Representation of a Preference Ordering by a Numerical Function," in *Decision Processes*, R.M. Thrall, C.H. Coombs, R.L. Davis, eds.. New York: Wiley, pp. 159–165.

DeMarzo, P.M. (1988): "Coalitions and Sustainable Social Norms in Repeated Games," mimeo, Stanford University.

Farrell, J. (1984): "Credible Repeated Game Equilibria," unpublished manuscript

Farrell, J., and E. Maskin (1987): "Renegotiation in Repeated Games," Working Paper 8759, Department of Economics, University of California, Berkeley.

Greenberg, J. (1988): "Social Situations: A Game–Theoretic Approach," forthcoming, Cambridge University Press, Cambridge.

Güth, W., W. Leininger, and G. Stephan (1988): "On Supergames and Folk Theorems: A Conceptual Discussion," ZiF–Working Paper: Game Theory in the Behavioral Sciences 1987–88, 19, University of Bielefeld.

Harsanyi, J.C., and R. Selten (1988): "A General Theory of Equilibrium Selection in Games," forthcoming.

Nash, J. (1950): "The Bargaining Problem." *Econometrica*, 18, pp. 155–162.

von Neumann, J., and O. Morgenstern (1947): Theory of Games and Economic Behavior, Princeton: Princeton University Press.

Pearce, D.G. (1987): "Renegotiation–Proof Equilibria: Collective Rationality and Intertemporal Cooperation," Cowles Foundation Discussion Paper No. 855.

Selten, R. (1965): "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit," *Zeitschrift für die gesamte Staatswissenschaft*, 12, pp. 301–324.

Selten, R. (1973): "A Simple Model of Imperfect Competition where 4 are Few and 6 are Many," *International Journal of Game Theory*, 3, pp. 141–201.

Selten, R. (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, pp. 22–55.