ECONOMIC RESEARCH REPORTS

Monotone Games with Positive Spillovers

by

Douglas Gale

RR# 98-34

November 1998

C.V. STARR CENTER FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Monotone Games with Positive Spillovers[1]

Douglas Gale
Department of Economics
New York University
269 Mercer Street
New York, NY 10003

May 1997; revised October 1998.
J.E.L Classification: C73, H41
Keywords: monotone games, positive spillovers, efficiency, subgame perfect equilibrium, voluntary contribution games.

[1]Earlier versions of this paper were presented in seminars at New York University and the London School of Economics and at the Northwestern University Summer Workshop (Microeconomics Week), where participants made many useful comments. I am particularly indebted to Bob Rosenthal for reading the paper and to Steven Matthews for pointing out the relevance of his own work. The financial support of the National Science Foundation, under grants SBR 9422194 and SBR 9810912, and the C. V. Starr Center for Applied Economics are gratefully acknowledged.

Abstract

A monotone game is a repeated game in which action sets are partially ordered and players' actions over time are non-decreasing with respect to the partial order. These games represent a variety of strategic situations in which players are able to commit to certain levels of action.

1 Introduction

Coordination games represent an important class of examples that show how individually rational behavior may lead to collectively suboptimal outcomes, even when economic agents have common interests. Game theorists are interested in these games because of their surprising departure from collective rationality. Macroeconomists are interested in these games because they provide a paradigm for macroeconomic market failures.

A familiar example of a coordination game is the 2×2 game depicted below.

	L	R
T	$(2, 2)$	$(0, 0)$
B	$(0, 0)$	$(1, 1)$

For each strategy profile, the two players' payoffs are identical so they have completely aligned interests. However, there are two, Pareto-ranked equilibria, (Top, Left) and (Bottom, Right). The fact that rational players can end up choosing (Bottom, Right) despite their common interest in choosing (Top,Left) suggests a surprising limitation of the power of rationality to guarantee satisfactory outcomes.

Now the existence of multiple, Pareto-ranked equilibria in the static game does not guarantee that a bad equilibrium will be played. In fact, it can be argued that the best (Pareto-dominant) equilibrium will be focal in the sense of Schelling and hence will be chosen by rational players. One way of making this kind argument more compelling is to cast the game in a dynamic framework in which one can explicitly analyze the players' attempt to "coordinate" on an efficient equilibrium of the coordination game.

Several studies have taken this approach. Auman and Sorin (1989) studied the repeated play of two-person coordination games when players had finite memory and were uncertain about the rationality of their opponents. They showed that the Pareto-dominant outcome would be played in the long run with high probability. Lagunoff and Matsui (1997) studied the repeated play of n-person coordination games under the assumption that the players could only change their strategies at times that were independently and exponentially distributed. Here too the Pareto-dominant outcome is observed with high probability in the long run. Gale (1995) studied a game of timing and delay in which players had to decide when to to exploit an investment option, the returns to which depended on the number of other players who chose to invest. This was a dynamic version of a static coordination game, but not a repeated game. There are many equilibria in this game, but Gale showed that when the period length became vanishingly short, all equilibrium outcomes converged to the unique Pareto-dominant outcome.

In all of these studies, the dynamic structure added to the static coordination game fundamentally changes the analysis and provides an explanation of why the Pareto-dominant equilibrium should be selected. The aim of the present paper is to extend and clarify the analysis in Gale (1995). In Gale's model, investment is irreversible and players can only invest once, so once a player has invested he is committed to that decision forever. At any stage of the dynamic game, the continuation subgames is completely characterized by the number of players who have already invested. So subgames are ordered in a natural way and we can analyze the dynamic game by using backward induction on the ordered family of subgames. The decision to invest transforms a subgame into the next subgame. Once enough players have invested, it becomes a dominant strategy for the remaining players to invest. Using backward induction, one can show that there is a limit on the delay that can exist between successive ordered subgames, since it will always be in someone's interest to precipitate the next subgame, anticipating that someone else will precipitate the next, ..., until investment becomes profitable.

The model in Gale (1995) is very special but the irreversibility structure, which is responsible for the optimality properties of the dynamic coordination game, has a much wider applicability. To study the role of irreversibility in general, a general class of games with a monotone structure is introduced in Section 1. These are called *monotone games*. A monotone game is defined by a one-shot, n-player game Γ together with an irreversibility structure. The strategy set of player i is denoted by X_i and is endowed with some partial ordering $\geq_i$. The ordering is extended in the natural way to the strategy profiles $X = \times_{i=1}^{n} X_i$. Then the game Γ is played repeatedly over time, subject to the following two constraints.

- *Sequential Moves:* At each date t only one of the players i is allowed to change his strategy. The other players retain their strategy from the previous period.

- *Monotonicity:* If player i is allowed to change his strategy at date t then he must choose $x_i' \geq_i x_i$, where x_i is his previous strategy.

Both restrictions are important. The fact that only one player moves at a time implies that monotone games are games of perfect information (compare

Lagunoff and Matsui (1997)). This obviously gives backward induction a bigger role, though it is not necessary in the game studied in Gale (1995). The monotonicity restriction is the essential irreversibility condition.

Even without the monotonicity assumption, the assumption that players move one at a time can ensure efficiency in certain special cases. In Section 2 we study the class of games with common preferences. A game Γ has common preferences if each player has a payoff function $u_i : X \to \mathbf{R}$ and $u_i = u_j$ for all $i, j = 1, ..., n$. Coordination games are a special case of games with common preferences.

With common preferences and one player moving at a time, a repeated game is like a dynamic programming problem. At any date, the player who is allowed to move chooses a strategy to maximize the common payoff anticipating that all future players will do the same. This sounds just like solving a dynamic program, so it is not too surprising that the outcome is optimal in the limit as the period length becomes vanishingly small. Nonetheless, it points to some important features of these games. The "folk theorem" fails to hold because of the lack of simultaneous moves. If there were simultaneous moves, every Nash equilibrium would be a subgame perfect equilibrium. Discounting is also necessary. If players are not at least slightly impatient, it is easy to provide counterexamples.

The games studied by Admati and Perry (1991) and Gale (1995) do not have common preferences. In the contribution game, every player would like the other players to make larger contributions and his own contribution to be reduced. In the investment game, no player wants to invest before the minimum aggregate investment needed to ensure profitability has been reached. Consequently, in an asymmetric equilibrium, in which some players invest before others, the early investors would like to switch places with the later investors. For these games, the irreversibility assumption plays an important role.

In Section 3, we look at monotone games that satisfy *positive spillovers*. A game Γ has (strictly) positive spillovers if an increase in one player's strategy (increases) does not decrease the payoffs of all the other players. This is true of the contribution game and the investment game. It is also true of many coordination games. Then we characterize the subgame perfect equilibrium outcomes of monotone games with positive spillovers.

Again it is possible to show that sequential moves and monotonicity imply optimality, at least in the sense that the set of subgame perfect equilibrium outcomes includes a large number of Pareto-efficient outcomes. These are not the only outcomes that can be supported by equilibrium play, but the fact that they can be supported at all with the limited punishments allowed by monotonicity is surprising.

The structure of these equilibria is quite different from those of repeated games. In particular, there is no analogue of the "folk theorem". Instead, the subgame perfect equilibrium outcomes are precisely the *approachable* outcomes, where an outcome is defined to be approachable if it is the limit of a non-decreasing sequence along which no player can ever guarantee himself more than the payoff he receives at the desired outcome. In Section 4 we provide sufficient conditions for an outcome to be approachable. An outcome is *strongly minimal* if it is impossible for any coalition to guarantee each of its members at least as high a payoff by choosing smaller strategies, while every member of the counter-coalition chooses the minimal strategy. This condition sounds a bit like the definition of the core of Γ and there is indeed a family resemblence. What this makes clear is that extreme distributions of utility will be ruled out as subgame perfect equilibria, even though they may be efficient. In particular, a player can typically be sure of doing better than his security level in any SPE, so the set of SPE payoffs in a monotone game is smaller than the SPE payoffs of a repeated game.

Characterizing the set of other SPE outcomes is more difficult, but we can at least say that the set of pure Nash equilibrium outcomes of the game Γ is always contained in the set of SPE outcomes. The Nash property immediately implies strong minimality in games with positive spillovers.

The results obtained so far obviously leave a great deal unsaid. This should not be surprising when we recall that some special structure is required to ensure that monotonicity restricts the set of equilibria. Furthermore, it is too much to hope that monotonicity always ensures efficient outcomes. So it will not be possible to deduce interesting characterization theorems for monotone games in general. Some restrictions are needed if we are to have any hope of obtaining positive results. On the other hand, the more economic structure we are prepared to build into these games, the more interesting results we may hope to derive. This preliminary study suggests that future research on monotone games may prove to be very fruitful indeed.

2 Monotone Games

2.1 The Game

A monotone game differs from an ordinary repeated game in two respects. First, the players' strategies are nondecreasing over time. Secondly, the players move one at a time according to an exogenously specified order. These restrictions transform what would have been a repeated game into a stochastic game. The state of the game, which uniquely defines the subgame, consists of the date and the profile of strategies that were chosen in the last period

The definition of the monotone game begins with an n-player stage game Γ consisting of a set of players

$$N = \{1, ..., n\},$$

a set of strategy profiles

$$X = X_1 \times ... \times X_n,$$

and a payoff function player i's payoff function is denoted by

$$u : X \to \mathbf{R}^n.$$

The payoff function $u(x) \equiv (u_1(x), ..., u_n(x))$ specifies a payoff $u_i(x)$ for each player i.

The monotone game is played at a countable number of dates $t = 1, 2, ...$ The *order of play* is a sequence $\pi = \{\pi_t\}_{t=1}^{\infty}$ that specifies the identity of the player $i = \pi_t$ who is to move at date t. For example, we could assume that players take turns according to their index i: first player 1 moves, then player 2 moves, and so on until all the players have moved once. Then they begin again with player 1, player 2, and so on indefinitely. In that case, π is defined by $\pi_t = t \bmod n$, for any t. The one restriction needed on the order of play is that every player moves sufficiently often: for some $\varepsilon > 0$ and any date $T \geq n$ and any player i,

$$\#\{1 \leq t \leq T | \pi_t = i\} \geq \varepsilon T.$$

The strategy sets X_i are assumed to be subsets of some finite-dimensional Euclidean space $\mathbf{R}^\ell$. Each strategy set is given the usual partial ordering and the set of strategy profiles is ordered in the obvious way,

$$x' \geq x \Longleftrightarrow x'_i \geq x_i, \forall i.$$

The set of strategy profiles is assumed to be *comprehensive* in the sense that for any $x, x' \in \mathbf{R}^{n\ell}$,

$$[x \in X \text{ and } x' \geq x] \Longrightarrow x' \in X.$$

There is assumed to be an exogenously given initial state $x^0 \in X$. Since the strategy profiles are comprehensive, by a simple translation of the strategy sets we can set the initial state we can assume without loss of generality that the initial state of the game is $x^0 = 0$ and the set of strategies available to player i is $X_i = \mathbf{R}^{\ell}_{+}$.

Let $M(\Gamma)$ denote the monotone game defined in this way.

2.2 Equilibrium

A feasible path for $M(\Gamma)$ is a non-decreasing sequence $\{x_t\}_{t=1}^{\infty}$ satisfying the order-of-play restrictions that for every i and t

$$i \neq \pi_t \Longrightarrow x_{it+1} = x_{it}.$$

The informations sets of the game consist of the finite initial segments of feasible paths. Let H be the set of information sets and let $H_i \subset H$ be the set of information sets at which player i controls the play. A strategy for player i in $M(\Gamma)$ is a function $f_i : H_i \to X$ with the property that for each information set $h \in H_i$ the following two conditions are satisfied:

$$\begin{aligned} \text{(i) } x_{t+1} &= f_i(h) \geq x_t \\ \text{(ii) } x_{jt+1} &= f_{ij}(h) = x_{jt}, \forall j \neq i, \end{aligned}$$

where $h = (x_1, ..., x_t)$. Let F_i denote the set of feasible strategies for player i and let $F = F_1 \times ... \times F_n$.

For any n-tuple of feasible strategies $f = (f_1, ..., f_n) \in F$ there is a unique equilibrium path $\{x_t\}$ defined recursively by putting

$$\begin{aligned} x_1 &= f_{\pi_1}(0) \\ x_{t+1} &= f_{\pi_t}(x_t), \forall t > 1. \end{aligned}$$

The payoff to player i is denoted by $U_i(f)$. The payoffs can be specified in various ways. There are two cases that are of particular interest, with and without discounting. Without discounting, payoffs are defined by

$$U_i(f) = \lim \inf_{t \to \infty} u_i(x_t).$$

In the case with discounting, we assume there is a common discount factor $0 < \beta < 1$ and the payoffs are defined by

$$U_i(f) = (1-\beta)\sum_{t=1}^{\infty} \beta^{t-1} u_i(x_t).$$

A monotone game is a game of complete information and perfect recall. Therefore we analyze the subgame perfect equilibria of the game. Let $M_t(\Gamma, x)$ denote the monotone game beginning at date t with initial state x, that is, the stage-game strategies are restricted to the set $\{x' \geq x\}$ and the order of play is $\{\pi_t, \pi_{t+1}, ...\}$. A *subgame perfect equilibrium* is a strategy profile $f \in F$ such that for any history $h = (x_1, ..., x_{t-1}) \in H$, the continuation strategies $f|h$ constitute a Nash equilibrium of $M_t(\Gamma, x_{t-1})$.

3 Games with Common Preferences

Lagunoff and Matsui (1997) study repeated coordination games under the assumption that players can only change their strategies at exogenous, random intervals. This assumption of inertia implies that whenever an individual player has the opportunity to change his strategy, he can assume that all the other players will hold their strategies constant for some considerable time. He also anticipates the effect that his choice will have on the next player to have an opportunity to change his strategy. In this context, subgame perfect equilibrium implies that eventually they must all coordinate on the Pareto dominant equilibrium.

The first example we consider is a game with *common preferences*. Formally, a stage game Γ has common preferences if $u_i = u$ for $i = 1, ..., n$. This is a slight generalization of the familiar *coordination game*, in which players have common preferences and, in addition, prefer the diagonal outcomes to all others. Even in the very special case of coordination games, there may be multiple equilibria, some of which are Pareto-preferred to others. From this it is clear that common preferences do not guarantee good outcomes. Nonetheless, for monotone games based on stage games with common preferences, the results are quite different: in the limit, as the discount rate becomes vanishingly small, all subgame perfect equilibrium outcomes are approximately Pareto-efficient.

The reason has nothing to do with the monotonicity restrictions, but simply the fact that sequential decision making by the players transforms a strategic game into a maximization problem. It should be noted that the limiting result does not hold in the limit, where there is no discounting. In this sense, a small amount of discounting acts as a refinement of the subgame perfect equilibrium.

Let $M_t(\Gamma, x, \beta)$ denote the monotone subgame based on the stage game Γ that begins at date t with initial state x. Note that the subgame differs from $M(\Gamma, x, \beta)$ because the order of play is given by $(\pi_t, \pi_{t+1}, ...)$. Let $v_t(x)$ denote the infimum of the payoffs from any subgame perfect equilibrium of the subgame $M_t(\Gamma, x)$. We assume that u is bounded on X and hence $v_t(x)$ is bounded for every t and x.

$$\begin{aligned} v_t(x) &\geq \sup\{(1-\beta)u(x') + \beta v_{t+1}(x') : x' \in S_t(x)\} \\ &= \sup\{(1-\beta)u(x') + \beta \sup\{(1-\beta)u(x'') + \beta v_{t+2}(x'') : x'' \in S_{t+1}(x')\} : x' \in S_t(x)\} \\ &= \sup\{(1-\beta)u(x') + (1-\beta)\beta u(x'') + \beta^2 v_{t+2}(x'') : x'' \in S_{t+1}(x'), x' \in S_t(x)\} \end{aligned}$$

and so forth, where $S_t(x) \equiv \{x' \in S(x) : x'_i = x_i, \forall i \neq \pi_t\}$ and $S(x) \equiv \{x' \in X | x' \geq x\}$.. From this we have

$$v_t(x) = (1-\beta)\sup\{\sum_{k=0}^{\infty} \beta^k u(x_{t+k}) : x_{t+k} \in S_{t+k}(x_{t+k-1}), \forall k; x_{t-1} = x\}.$$

If u possesses a maximum on X, say x^*, and $x \leq x^*$, then there exists a sequence $\{x_{t+k}\}_{k=0}^{\infty}$ such that $x_{t-1} = x$, $x_{t+k} \in S_{t+k}(x_{t+k-1})$ for $k = 0, 1, ...$, and $x_t = x^*$ for all but a fixed, finite number of dates. From this it follows that $\lim_{\beta \to 1} v_t(x) = u(x^*)$ for all t and x^*.

Theorem 1 *(Lagunoff-Matsui) Let $\Gamma = (N, X, u)$ be a stage game with common preferences and suppose that u possesses a maximum at x^*. Then for any initial state $x \leq x^*$ and date t,*

$$\lim_{\beta \to 1} v_t(x; \beta) = u(x^*),$$

where $v_t(x; \beta)$ is the infimum of the payoffs from the subgame perfect equilibria of the monotone game $M_t(\Gamma, x, \beta)$.

Although the proof respects the monotonicity conditions of the game, they are not used at any point to obtain the result. So the theorem holds for any dynamic game of this form, even without the constraint that the states must be non-decreasing over time. This shows that it is only the sequential decision making, together with a positive but vanishing amount of discounting, that ensures optimality in the limit.

An example shows that the result fails if there is no discounting. Suppose there are two players $i = 1, 2$ with strategy sets $X_i = [0, 2]$ and payoffs

$$u(x) = -(\max\{x_1, x_2) - 1)^2.$$

Clearly the maximum of $u(x)$ is 0 and yet without discounting the outcome $x = (2, 2)$ can be supported as a subgame perfect equilibrium. To see this, consider the strategies that require each player to set $x_i = 2$, if he has not already done so, at each date when he has the chance to move. Each player, anticipating that his opponent will choose the maximum action and ensure a future payoff of -1 will find it optimal to follow the strategy described.

Discounting makes a difference to this analysis, because even if the player anticipates a poor outcome in the future, he would still like to postpone it as long as possible. If both players do this, they end up cooperating indefinitely.

[The common preference game is much simpler than the Lagunoff-Matsui game, but there is a strong family resemblence between the two. The game analyzed by Gale (1995) is quite different, because players do not have common preferences. The n players each have to decide when to exploit an investment option. The size of the investment is fixed and the investment can only be made once so effectively the players are deciding to invest at date t, where $t = \infty$ corresponds to waiting forever. Investment involves a fixed cost c incurred at the date the investment is made and a stream of revenues after the investment is made. The revenue in each period depends on the number of of players who have invested to date. Investment is assumed to be socially desirable and players are impatient, so the unique Pareto-efficient outcome is the one in which all players invest at the first date.

There are many equilibria and, since players discount the future, any delay in investment is inefficient. However, as the period length becomes vanishingly short (the discount factor converges to 1), all subgame perfect equilibrium outcomes converge to the Pareto-efficient outcome.

Because of the irreversibility of investment, subgames can be grouped into equivalence classes of strategically equivalent games according to the number of players who have already invested. There is a natural ordering on these equivalence classes, since subgames with more invested players follow those with less. If the number of invested players is large enough, it is a dominant strategy for all to invest. The entire game can be analyzed by backward induction, starting with the dominance solvable subgames and working backwards to the preceding subgames. This backward induction argument can be used to extend the analysis of games with common preferences to games with dominance solvable subgames, but monotonicity plays a role here that is not found in the games with common preferences.]

4 Games with Positive Spillovers

In this section we examine monotone games in which each player's payoff is monotonically non-decreasing in the actions of other players. Following Cooper and John (1988), we say that a stage game Γ exhibits *positive spillovers* if, for any player i and any $x_{-i}, x'_{-i} \in X_{-i}$,

$$u_i(x_i, x_{-i}) \geq u_i(x_i, x'_{-i}) \text{ if } x_{-i} \geq x'_{-i}, \text{ for any } x_i \in X_i.$$

Analogously, Γ is said to exhibit *strictly positive spillovers* if, for any player i and any $x_{-i}, x'_{-i} \in X_{-i}$,

$$u_i(x_i, x_{-i}) > u_i(x_i, x'_{-i}) \text{ if } x_{-i} > x'_{-i}, \text{ for any } x_i \in X_i.$$

The assumption of positive spillovers is maintained throughout the rest of this section. The stage game Γ is also assumed to satisfy the following regularity assumptions:

- The set of individually rational actions for player i is bounded: for some large number k and any $x \in X$,

$$u_i(x_i, x_{-i}) \leq u_i(0) \text{ if } \|x_i\| > k.$$

- For each player i, the payoff function u_i is assumed to be continuous on $\mathbf{R}^{\ell}_{+}$.

Example 2 *An example of a game with strictly positive spillovers is the* public goods contribution game*. Each of the players $i = 1, ..., n$ contributes an amount $x_i \geq 0$ of the numeraire good for the provision of a public good. The total amount of the public good produced is $y = f(\sum_i x_i)$, where $f(\cdot)$ is a continuous and increasing production function. Each player i has a utility function $W_i(y, m_i)$, where y is player i's consumption of the public good and $m_i = \bar{m}_i - x_i$ is his consumption of the numeraire good. $W_i(\cdot)$ is assumed* *to be continuous and increasing in both arguments. Then the payoff function can be defined by putting*

$$u_i(x) = W_i(y, \bar{m}_i - x_i),$$

where $y = f(\sum_j x_j)$*. It is clear that the payoff function* $u_i(\cdot)$ *defined in this way is continuous and exhibits strictly positive spillovers.*

Admati and Perry (1991) study a similar game in which two players make successive contributions to the production of a public good. The public good is assumed to be socially desirable and has a fixed cost K. The players make contributions alternately, beginning with player 1 in period 1. The contributions are sunk (non-recoverable) and players incur the cost of each contribution at the date it is made. The public good is produced and the players enjoy the benefit of it as soon as the contributions total K. Players are impatient and discount the future using a common discount factor $0 < \delta < 1$. Admati and Perry show that under certain conditions there is an essentially unique equilibrium. The equilibrium may be inefficient: in the case of linear costs, the project is completed if and only if each player would find it optimal to complete the project if he were the only player. This condition is restrictive. The lumpiness of the project is crucial here. As we shall see, if the public good is perfectly divisible, it may be possible to support Pareto-efficient equilibria under weaker conditions.

Another source of inefficiency in the Admati-Perry model comes from delay in completing the project: since players discount the future, there is a loss of utility whenever the project is not completed at the earliest possible date. However, in the case where the project is completed, it can be shown that as the period length converges to zero (the discount factor converges to 1) the unique subgame perfect equilibrium outcome converges to the Pareto-dominant outcome. In other words, there is virtually no delay.

Marx and Matthews (1997) extend the model of Admati and Perry (1991) and provide a thorough investigation of a number of interesting points. In particular, they suggest that simultaneous moves may improve efficiency by admitting equilibria in which the public project can be provided even though neither of the players would be willing to provide it individually. In a simultaneous move game, each player is in effect willing to provide part of the cost on the assumption (validated in equilibrium) that the other player is providing the rest. This raises the question of equilibrium robustness: if players are unable to coordinate their moves perfectly, the efficient equilibrium may disappear, leaving only the sequential move equilibrium.

In a game with positive spillovers, the worst threat that players $j \neq i$ can make against player i is to maintain their current actions. Suppose that the current state of the game is $x \in X$. Player i's reservation level at the current state x is defined to be the best payoff that player i can guarantee by his own actions in any single period. The reservation level is denoted by $u_i^*(x)$ and defined by putting

$$u_i^*(x) = \sup_{x_i' \in X_i} \{u_i(x_i', x_{-i}) | x_i' \geq x_i\}.$$

We note without proof that u_i^* satisfies the same properties as u_i.

Lemma 3 $u_i^* : X \to \mathbf{R}$ *is a well defined, continuous function. It is non-increasing in* x_i *and non-decreasing in* x_j *for* $j \neq i$*. If* Γ *exhibits strictly positive spillovers, then* u_i^* *is increasing in* x_j *for all* $j \neq i$*.*

Having reached a state x player i can always guarantee himself at least $u_i^*(x)$ in the continuation game, so in the long run u_i^* is a better measure of his welfare in state x. In the sequel, we use u_i^* in place of the payoff function u_i in analyzing the game $M(\Gamma)$.

A *satiation point* is a state $x \in X$ such that $u^*(x) = u(x)$. Satiation points play a role somewhat like Nash equilibria in the analysis of monotone games. Once a satiation point has been reached, no player can make himself better off by unilaterally deviating. Hence it is a subgame perfect equilibrium to remain at a satiation point forever.

Let x^0 be the status quo and let $x \in S(x^0) \equiv \{x \in X | x \geq x^0\}$. The state x is *approachable* in $S(x^0)$ if there exists a feasible path $\{x^t\}_{t=1}^{\infty}$ starting at x^0 such that (a) $u_i^*(x^t) \leq u_i(x)$, for every i and t, and (b) $x^t \to x$ as $t \to \infty$.

Note that if x is approachable in $S(x^0)$, then the continuity of u^* implies that $u^*(x) \leq u(x)$. Since $u^*(x) \geq u(x)$, it follows that any approachable state is a satiation point.

An *equilibrium outcome* of the game $M_t(\Gamma, x)$ is the limit of an equilibrium path for some SPE of $M_t(\Gamma, x)$. The main result of this section is to show that the approachable states are precisely the equilibrium outcomes of a monotone game with positive spillovers.

Theorem 4 *Under the maintained assumptions, for any initial state* x*, a state* $x' \in S(x)$ *is an equilibrium outcome of* $M_t(\Gamma, x)$ *if and only if* x' *is approachable in* $S(x)$*.*

Before attempting the proof of the theorem, it will be helpful to prove the following simpler result.

Lemma 5 *Under the maintained assumptions, for any state x there exists a state $x' \in S(x)$ that is approachable in $S(x)$.*

Proof. The proof is constructive. Choose a feasible sequence of states $\{x_t\}_{t=1}^{\infty}$ to satisfy the following conditions:

$$x_t \in \arg\max\{u_{\pi(t)}(x') : x' \in S_{\pi(t)}(x_{t-1})\}.$$

In other words, at each date t let the player $\pi(t)$ who controls the play at that date choose a myopic best response to the current state, subject to the condition that states be non-decreasing. The feasible path $\{x_t\}_{t=1}^{\infty}$ is non-decreasing by construction and bounded because $u(x_t) \geq u(0)$ requires $\|x_{it}\| \leq k$ for some k and all (i,t). So there exists a limit point x^* such that $x_t \to x^*$. Also, $\{u(x_t)\}_{t=1}^{\infty}$ is non-decreasing because there are positive spillovers and the player who moves chooses a best response. That is, $u_i(x_t) \geq u_i(x_{t-1})$ for $i \neq \pi(t)$ because $x_t \geq x_{t-1}$ and $x_{it} = x_{it-1}$ and $u_i(x_t) = u_i^*(x_{t-1}) \geq u_i(x_{t-1})$ for $i = \pi(t)$. Since $\{u(x_t)\}_{t=1}^{\infty}$ is non-decreasing and $u_i^*(x_t) = u_i(x_{t-1})$ infinitely often, for each i, it follows that $u^*(x_t) \leq u(x^*)$ for all t. In other words, if $u_i^*(x_t) > u_i(x^*)$ for some i and t, there would be a later date $t' > t$ at which $i = \pi(t')$ gets to move and then

$$u_i(x_{t'}) = u_i^*(x_{t'-1}) \geq u_i^*(x_t) > u_i(x^*),$$

a contradiction of what has already been proved. This completes the proof that x^* is approachable in $S(x)$. ■

The importance of this lemma is that it allows us to show that certain states that we want to use as punishments are also approachable.

Lemma 6 *Under the maintained assumptions, for any state x and any player i there exists a state $x' \in S(x)$ that is approachable in $S(x)$ and satisfies $u_i(x') = u_i^*(x')$. In other words, we can use x' to hold i's payoff down to his reservation level in the subgame $M(\Gamma, x)$.*

Proof. For any initial state x and any $x_i' \in S_i(x)$, let $\zeta(x_i')$ denote the (non-empty) set of approachable outcomes in $S(x_{-i}, x_i')$. It is easy to see that $\zeta : X \to X$ is upper hemi-continuous and that $\phi(x_i') = \sup\{u_i(x') :$

$x' \in \zeta(x_i')\}$ is upper semi-continuous. The set $\{x_i' \in S_i(x) : \phi(x) \geq u_i^*(x)\}$ is non-empty because $\phi(x) \geq u_i^*(x)$ by construction; bounded because the set of individually rational states is bounded; and closed because ϕ is upper semi-continuous. There exists a maximal element x_i^0 in this compact set and it clearly satisfies $\phi(x_i^0) = u_i^*(x)$. To see this, suppose to the contrary that $\phi(x_i^0) > u_i^*(x)$. Then by u.s.c. there exists $x_i' > x_i^0$ such that $\phi(x_i') > u_i^*(x)$, a contradiction. We have now shown that there exists an approachable state x^1 in $S(x_{-i}, x_i^0)$ such that $u(x^1) = u_i^*(x)$.

Now we can show that x^1 is approachable in $S(x)$ as follows. Construct a feasible path $\{x_t\}_{t=1}^{\infty}$ by putting

$$x_t = x, t = 1, ..., k-1$$

where k is the first date at which $i = \pi(t)$, putting

$$x_k = (x_{-i}, x_i^0)$$

and putting

$$x_t = x_{t-k}'$$

for $t > k$, where $\{x_t'\}_{t=1}^{\infty}$ is the feasible path by which x^1 is approachable in $S(x_{-i}, x_i^0)$. The path so constructed has all the required properties, so x^1 is approachable in $S(x)$. ■

Now we are ready to prove the theorem. Suppose that x is an equilibrium outcome. Then there exists an equilibrium path $\{x_t\}$ converging to x. To show that x is approachable, we only need to show that $u_i^*(x_t) \leq u_i(x)$ for every t and i. Suppose to the contrary that $u_i^*(x_t) > u_i(x)$. Then at the next date t' when i gets to move, $u_i^*(x_{t'}) > u_i(x)$ since $\{x_t\}$ is non-decreasing and so is u_i^*. Then at this date player i can do better than his equilibrium payoff since he can attain a state x' such that $u_i(x') > u_i(x) = U_i(f) = \liminf_{t\to\infty} u_i(x^t)$. Whatever happens after this his payoff cannot be lower than $u_i(x')$ since other players cannot reduce their strategies and there are positive spillovers.

To prove the converse we assume that x is an approachable state and construct a SPE. The equilibrium path consists of a feasible path $\{x_t\}$ converging to x. Suppose that at some date t, player i deviates to a state x_t'. From that moment on the play switches to some target state x^i, say, which is approachable from x_t' and yields a payoff to i equal to $u_i^*(x_t') \leq u_i^*(x_{t-1}) \leq u_i(x)$. In other words, the play follows a feasible path approaching x^i. If there is a further deviation, it is dealt with in a similar way. A complete strategy for all players is constructed in this way and the result is clearly a SPE since no deviation will ever earn the player more than his reservation utility, which in turn does not exceed the utility of x. ■

5 Approachability

Although approachability provides us with what we need to characterize subgame-perfect equilibrium outcomes, it does not provide us with a criterion stated in terms of the primitives of the model. In this section, sufficient conditions that are at least more closely related to the primitives of the model are explored.

A state x is said to be *locally approachable* if, for any $\varepsilon > 0$, there exists another feasible state $\bar{x} \leq x$ within a distance ε of x and possessing the properties that $u(\bar{x}) \leq u(x)$ and $u_i(\bar{x}) < u_i(x)$ for some i. A state x is *positive* if $x_i > 0$ for all i and *minimal* if it is positive and there does not exist a state $x' < x$ such that $u^*(x') \geq u^*(x)$.

Lemma 7 *If x^* is a minimal satiation point, then x^* is locally approachable.*

Proof. Fix some small number $0 < \varepsilon < 1$ and each i define x^i by putting

$$x_j^i = \begin{cases} (1-\varepsilon)x_i^* & \text{if } j = i \\ x_j^* & \text{if } j \neq i. \end{cases}$$

Since x^* is minimal and $x^i < x^*$, for each i, $u_j(x^i) < u_j(x^*)$ for some j.

For fixed ε, let

$$K = \text{co}\{x^i : i = 1, ..., n\}$$

and let

$$A_i = \{x \in K |\, u_i^*(x) \geq u_i^*(x^*)\}.$$

Since u^* is continuous and K is closed, A_i is closed for every i. The fact that x^* is minimal implies that $\cap_i A_i = \emptyset$.

For any $x \in K$ let $g_i(x)$ denote the distance between x and A_i, that is,

$$g_i(x) = \min_{x' \in A_i} \|x - x'\|.$$

Without loss of generality we can identify K with the unit simplex $\{t \geq 0 | \sum_i t_i = 1\}$ and then we can define a mapping T from K to itself by putting

$$T_i(x) = \frac{g_i(x)}{\sum_j g_j(x)}.$$

Clearly, T is continuous because g_i is continuous and T is well defined because $\cap_i A_i = \emptyset$ implies that $\sum_j g_j(x) > 0$ for every $x \in K$. The conditions of Brouwer's fixed point theorem are satisfied, so there exists a point $\bar{x} \in K$ such that $T(\bar{x}) = \bar{x}$. Let $I = \{i = 1, ..., n | \bar{x}_i > 0\} = \{i = 1, ..., n | g_i(x) > 0\}$. Recalling that $g_i(x) > 0$ implies that $x \notin A_i$, we have that $u_i^*(\bar{x}) < u_i^*(x^*)$ for all $i \in I$. For all $i \notin I$ we have $\bar{x} < x^*$ and $\bar{x}_i = x_i^*$ so positive spillovers implies that $u_i^*(\bar{x}) \leq u_i^*(x^*)$. Then $\bar{x}$ is the required state and since ε is arbitrary, x^* is locally approachable. ■

The role of the property of local approachability is to ensure that we can get some room to maneuver near to the target state x^*. By moving to x^*, at least one of the constraints $u_i^*(\bar{x}) \leq u_i^*(x^*)$ is strictly satisfied and so we can reduce $\bar{x}_i$ a small amount without violating any of the other constraints. The next set of properties ensures that this process can be continued without getting "stuck".

A positive satiation point x^* is said to be *strongly minimal* if there does not exist a state $0 < x < x^*$ such that for every i

$$x_i > 0 \Rightarrow u_i^*(x) \geq u_i^*(x^*).$$

Strong minimality obviously implies minimality.

Lemma 8 *If x^* is a strongly minimal satiation point, then x^* is approachable in $S(0)$.*

Proof. From the previous lemma, x^* is clearly locally approachable, so for any small $\varepsilon > 0$ there exists $\bar{x} < x^*$ such that $\bar{x}$ is ε-close to x^* and $u^*(\bar{x}) < u^*(x^*)$.

Construct a sequence $\{x^k\}$in the following way. Choose $x^1 = \bar{x}$ and for each $k = 1, 2, ...$ choose x^{k+1} so that:

(a) $x^k - x^{k+1} \in Z_i, \exists i$
(b) $x^{k+1} \leq x^k$
(c) $u_i^*(x^{k+1}) \leq u_i^*(x^*), \forall i$

and (d) $\|x^k - x^{k+1}\|$ is a maximum subject to the constraints (a)-(c).

Clearly, there is a (possibly non-unique) way of continuing the sequence at each step. Furthermore, as long as $x^k \neq 0$, we must have $x^{k+1} \neq x^k$ because the fact that x is strongly minimal implies that $u_i^*(x^k) < u_i^*(x^*)$ for some i such that $x_i^k > 0$ for each k. Then the sequence x^k must reach 0 in a finite number of steps. If not, we can continue the sequence indefinitely and, by compactness, there exists a subsequence converging to a limit point x^∞, say. With a slight abuse of notation, let $\{x^k\}$ denote the subsequence that converges to x^∞. Since $\|x^k - x^{k+1}\| \to 0$ along this subsequence, in the limit we have $u^*(x^\infty) = u^*(x^*)$, contradicting the minimality of x. This contradiction implies that $x^k = 0$ for some finite k.

What we have shown so far is that for any choice of $\bar{x}$ we have a non-decreasing sequence $\{x^k\}_{k=1}^K$ with $x^K = \bar{x}$ and $u^*(x^k) \leq u(x)$ for every k. (Note that the ordering of the sequence has been reversed). Now suppose that we take a sequence $\{\bar{x}^n\}_{n=1}^\infty$ of target states converging to x and denote the corresponding approaching sequences by $\{\{x^{nk}\}_{k=1}^{K_n}\}_{n=1}^\infty$. Using the standard diagonalization procedure, we can select a sequence $\{x^\ell\}_{\ell=1}^\infty$ such that x^ℓ is a limit point of $\{x^{n\ell}\}$, for each fixed ℓ and $x^\ell \to x$ as $\ell \to \infty$. Then $\{x^\ell\}$ is the required sequence. ■

5.1 A Public Goods Example

To illustrate the idea of restrictions implied by the concept of approachability, we can consider an example of the public goods contribution game introduced earlier. The production technology for public goods exhibits constant returns to scale: one unit of labor produces one unit of the public good. There are n players and player i's payoff function is

$$u_i(x) \equiv a \sum_{j=1}^{n} x_j - \frac{1}{2}(x_i)^2,$$

where x_i is player i's production of the public good and $\sum_i x_i$ is the total production of the public good.

Since the payoff functions are concave, the Pareto-efficient outcomes of the game can be characterized as solutions to a problem of maximizing the weighted sum of the individual players' payoffs. An admissible set of weights is denoted by the vector $\lambda = (\lambda_1, ..., \lambda_n)$, where $\lambda_i > 0$ for each i and $\sum_i \lambda_i = 1$. We can ignore the possibility that $\lambda_i = 0$ because that would imply

$x_i = \infty$. Then an outcome $x = (x_1, ..., x_n)$ is Pareto-efficient if and only if it solves the following problem:

$$\max_x \sum_{i=1}^{n} \lambda_i \left(a \sum_{j=1}^{n} x_j - \frac{1}{2}(x_j)^2 \right). \tag{1}$$

for some positive set of weights λ. The first-order conditions for the maximization problem are necessary and sufficient:

$$a \sum_{j=1}^{n} \lambda_j - \lambda_i x_i \leq 0 \text{ and } \left(a \sum_{j=1}^{n} \lambda_j - \lambda_i x_i \right) x_i = 0, \ \forall i.$$

Since $\sum_{j=1}^{n} \lambda_j = 1$, the first-order conditions are equivalent to $x_i = a/\lambda_i$ for each i. Since the first-order conditions are necessary and sufficient, an outcome x is Pareto efficient if and only if it satisfies

$$x_i = a/\lambda_i, \forall i,$$

for some positive weights summing to one.

For any admissible set of weights $\lambda = (\lambda_1, ..., \lambda_n)$, the utility of player i at the solution to (1) can be written as

$$u_i^*(\lambda) = \sum_{j=1}^{n} \frac{a^2}{\lambda_j} - \frac{1}{2}\left(\frac{a}{\lambda_i}\right)^2$$

The same general formula applies to any coalition of players, if we make the appropriate adjustments for the number of players, on the assumption that the counter coalition contributes nothing. A coalition is any non-empty set of players $S \subseteq N \equiv \{1, ..., n\}$ and an admissible set of weights for this coalition is a vector λ such that $\lambda_i > 0$ for all i and $\sum_{i \in S} \lambda_i = 1$. The counter coalition consists of players in $N \backslash S$ and we assume an outcome is attainable for S if $x_i = 0$ for all $i \in N \backslash S$. An outcome is efficient for S if it is attainable for S and there is no other attainable outcome for S that makes some members of S better off without making other members of S worse off. By the earlier reasoning, an outcome x is efficient for S if and only if it maximizes a weighted sum of payoffs of players in S for some admissible set of weights. Suppose that λ is an admissible set of weights. Then the efficient payoff for player i in the coalition S is

$$u_i^*(\lambda; S) = \sum_{j \in S} \frac{a^2}{\lambda_j} - \frac{1}{2}\left(\frac{a}{\lambda_i}\right)^2.$$

In particular, for a singleton coalition $\{i\}$ there is a unique efficient outcome, which is the result of maximizing $ax_i - \frac{1}{2}(x_i)^2$. The solution is to put $x_i = a$ and the resulting payoff is $a^2 - \frac{1}{2}a^2 = \frac{1}{2}a^2$. This is the individual player's security level and any equilibrium outcome must give him at least this much.

A coalition S can *improve on* an outcome x if there exists an admissible set of weights λ for S such that $u_i^*(\lambda; S) > u_i(x)$ for every $i \in S$. A feasible outcome x belongs to the *core* of the game if there does not a coalition that can improve on x.

The core is non-empty because it contains the *equal treatment outcome*, the outcome that gives each player the same payoff. Putting $\lambda_i = 1/n$ for each i, the equal treatment outcome yields a utility of

$$\begin{aligned} u_i^*(1/n) &= \sum_{j=1}^{n} a^2 n - \frac{1}{2}a^2 n^2 \\ &= \frac{1}{2}a^2 n^2. \end{aligned}$$

For any m-member coalition S, the equal treatment outcome with no contributions by the counter coalition would yield a smaller payoff

$$u_i^*(1/m; S) = \frac{1}{2}a^2 m^2,$$

for all $i \in S$. For any other S-admissible weights λ,

$$\begin{aligned} u_i^*(\lambda; S) &= \sum_{j\in S} \frac{a^2}{\lambda_j} - \frac{1}{2}\left(\frac{a}{\lambda_i}\right)^2 \\ &\leq a^2\left(\sum_{j\in S} \max_{j\in S}\{\frac{1}{\lambda_j}\} - \frac{1}{2}\left(\frac{1}{\lambda_i}\right)^2\right) \\ &= a^2\left(m \max_{j\in S}\{\frac{1}{\lambda_j}\} - \frac{1}{2}\left(\frac{1}{\lambda_i}\right)^2\right). \end{aligned}$$

Now suppose that $\lambda_k = \min_{j\in S}\{\lambda_j\}$. Then

$$\begin{aligned} u_k^*(\lambda; S) &= a^2\left(\frac{m}{\lambda_k} - \frac{1}{2}\left(\frac{1}{\lambda_k}\right)^2\right) \\ &\leq a^2(m^2 - \frac{1}{2}m^2) = \frac{1}{2}a^2 m^2. \end{aligned}$$

So for any other weights, at least one member of the coalition does no better than the equal treatment outcome for the coalition S, which is strictly worse than the equal treatment outcome for the grand coalition N. Thus, no coalition can improve on the equal treatment outcome so the equal treatment outcome belongs to the core.

Any core outcome is clearly strongly minimal and hence approachable. In general, the core property is stronger than strong minimality because strong minimality requires that an "improvement" be made by reducing the strategies of the coalition members. In this example, we are interested in the efficient outcomes and it turns out that if an efficient allocation can be improved on by a coalition, then it can be improved on by reducing the strategies of the coalition members. So the efficient outcomes that are *not* approachable are not core outcomes either.

It is easy to see that not all efficient outcomes are approachable. We have already noted that an approachable outcome must be individually rational and this implies that player i cannot receive a payoff lower than $\frac{1}{2}a^2$.

In repeated games, the individual rationality constraint is the only one that needs to be satisfied. The Folk Theorem tells us that any individually rational payoff vector can be supported by a subgame perfect equilibrium. By comparison with folk-theorem type results, the set of approachable and hence equilibrium outcomes is smaller in monotone games. Take, for example, $n = 3$ and consider outcome corresponding to the weights $\lambda_1 = \lambda_2 = \varepsilon$ and $\lambda_3 = 1-2\varepsilon$. A single player can achieve $\frac{1}{2}a^2$, so individual rationality requires that

$$u_i^*(\varepsilon, \varepsilon, 1-2\varepsilon) \geq \frac{1}{2}a^2.$$

However, a coalition of players 1 and 2 can do even better. With no help from player 3, they can achieve $u_i^*(1/2; S) = \frac{1}{2}a^2 2^2 = 2a^2$, so coalition $\{1, 2\}$ imposes the constraint

$$u_i^*(\varepsilon, \varepsilon, 1-2\varepsilon) \geq 2a^2$$

or $\varepsilon \geq \frac{1}{4}$.

REFERENCES

Admati, Anat and Motty Perry. "Joint Projects without Commitment," *Review of Economic Studies* **58** (1991) pages 259-76.

Aumann, Robert and Sylvain Sorin. "Cooperation and Bounded Recall," *Games and Economic Behavior* **1** (1989) pages 5-39.

Cooper, Russell and Andrew John. "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics* **103** (1988) pages 441-63.

Gale, Douglas. "Dynamic Coordination Games," *Economic Theory* **5** (1995) pages 1-18.

Lagunoff, Roger and Akihiko Matsui. "An 'Anti-Folk Theorem' for a Class of Asynchronously Repeated Games," *Econometrica* **65** (1997) 1467-77.

Marx, Leslie and Steven Matthews. "Dynamic Voluntary Contribution to a Public Project," Northwestern University (1997) unpublished.