W91-5

6/20/91

# COVENANTS WITH AND WITHOUT A SWORD: SELF-ENFORCEMENT IS POSSIBLE

by

**Elinor Ostrom**
**Department of Political Science**
**and**
**Workshop in Political Theory and Policy Analysis**

AND

**James Walker and Roy Gardner**
**Department of Economics**
**and**
**Workshop in Political Theory and Policy Analysis**

Presented at the International Association for the Study of Common Property, Second Annual Meeting, Winnipeg, Canada, September 26-29, 1991.

An earlier version of this paper was presented at the Midwest Political Science Association Meetings, Chicago, Illinois, April 18-20, 1991; and at the 1991 Annual Meeting of the American Political Science Association, The Washington Hilton, Washington, D.C., August 29-September 1, 1991.  We would like to thank William Bianco, Werner Guth, and Rick Wilson for useful comments. Financial support from the National Science Foundation (Grant SES-4843901) is gratefully acknowledged. All data are stored on permanent NOVANET disk files. Send inquiries to Professor James M. Walker, Department of Economics, Ballantine 901, Indiana University, Bloomington, Indiana, 47405.

Workshop in Political Theory and Policy Analysis

Indiana University 513 North Park Bloomington, Indiana 47405-3186

**COVENANTS WITH AND WITHOUT A SWORD: SELF-ENFORCEMENT IS POSSIBLE**


## ABSTRACT

Contemporary political theory often assumes that individuals cannot make credible commitments where substantial temptations exist to break them, unless such commitments are enforced by an external agent. Empirical evidence suggests, however, that individuals facing social dilemmas in many cases develop credible commitments without relying on external authorities. Fishers, irrigators, or herders appropriating from a CPR have repeatedly shown their capacity to organize themselves, to establish credible commitments, to monitor each others' behavior, and to impose sanctions on those who break their commitments. In this paper, we present findings from a series of experiments designed to explore the issue of endogenous formation of commitments and enforcement of such commitments. In a laboratory environment designed to parallel the decision environment of many CPRs, we manipulate treatments to examine: (1) communication alone (one-shot and repeated), (2) sanctioning alone, and (3) communication combined with the possibility of sanctioning.

# COVENANTS WITH AND WITHOUT A SWORD: SELF-ENFORCEMENT IS POSSIBLE

> And Covenants, without the sword, are but words, and of no strength to secure a man at all (Hobbes 1960: 109).

## I. Introduction

Contemporary political theories are frequently built upon the presumption that individuals cannot make credible *ex ante* commitments where substantial *ex post* temptations exist to break them, unless such commitments are enforced by an external agent. Hobbes justified the necessity of Leviathan on the frailty of mere words. For Hobbes, a contract that involves a promise by at least one of the parties to perform in the future is called a "covenant" (Hobbes 1960: 87). When both parties promise future performance, it is a "covenant of mutual trust" (ibid., 89). A covenant of mutual trust in a state of nature is void in Hobbes's view if either has a reasonable suspicion that the other will not perform.

> For he that performeth first, has no assurance the other will perform after; because the bonds of words are too weak to bridle men's ambition, avarice, anger, and other passions, without the fear of some coercive power; which in the condition of mere nature, where all men are equal, and judges of the justness of their own fears, cannot possibly be supposed (ibid., 89-90).

On the other hand, a covenant made "where there is a power set up to constrain those that would otherwise violate their faith" is likely to be fulfilled (ibid.). Thus, Hobbes argued for the necessity of a "coercive power, to compel men equally to the performance of their covenants, by the terror of some punishment, greater than the benefit they expect by the breach of their covenant" (ibid., 94).

The weakness of mere words and the necessity of external agents to enforce contracts is also a foundation upon which the powerful edifice of noncooperative game theory has been constructed. John Nash (1950, 1951) distinguished between cooperative and noncooperative game theory using the dual criteria that players could not communicate freely nor make enforceable agreements in noncooperative games. Most contemporary game theorists rely primarily on the second criterion. Harsanyi and Selten, for example, stress that:

> the decisive question is whether the players can make enforceable agreements, and it makes little difference whether they are allowed to talk to each other. Even if they are free to talk and to negotiate an agreement, this fact will be of no real help if the agreement has little chance of being kept. An ability to negotiate agreements is useful only if the rules of the game make such agreement binding and enforceable (1988: 3).[1]

Thus, communication without a change in the payoff function cannot upset a Nash equilibrium.

The predicted noncooperative outcomes of a large family of social dilemma games have been of considerable puzzle to many theorists and policy analysts.[2] All social dilemma games share the following characteristics:

a. all equilibria are deficient;

b. there exist substantial gains to coordination of behavior;

c. norms are not in place for coordinating such behavior.[3]

When the participants in a social dilemma share full information about the structure of the game, the payoffs to each player, and know that every other player also has this information, the game is one of full information. In a finitely-repeated, full-information dilemma, equilibrium payoffs are persistently lower than if the participants could make credible *ex ante* commitments. The Hobbesian state of

---

**[1]Hananyi and Selten (1988: 3) add that in real life, "agreements may be enforced externally by courts of law, government agencies, or pressure from public opinion; they may be enforced internally by the fact mat the players are simply unwilling to violate agreements on moral grounds and know that this is the case." To model self-commitment using noncooperative game theory, the capability of breaking the commitment is removed by trimming the branches that emanate from a self-commitment move to remove any alternative contrary to that which has been committed.**

**[2]The most famous social dilemma game is the Prisoners' Dilemma (PD). We have purposely not limited our discussion to the PD, as the class of social dilemmas is much larger.**

**[3]A social dilemma is only a *dilemma* from the perspective of noncooperative game theory. Within the context of cooperative game theory, the presumption is made that individuals will find a way to make a commitment to one another even though they will later face substantial temptations to break that commitment.**

nature has frequently been represented as a social dilemma, illustrating the close affinity between Hobbes's analysis and noncooperative game theory.[4]

The dual assumptions that "the sword" and "an external enforcer" are necessary before individuals can make credible *ex ante* commitments in social dilemma situations have important implications.[9] Many environmental problems-including those related to common-pool resources (CPRs)-can be represented as social dilemmas. As soon as this step is taken, policy recommendations of a particular type follow almost immediately. Ophuls (1973: 228) argued, for example, that "environmental problems cannot be solved through cooperation . . . and the rationale for government with major coercive powers is overwhelming." His conclusion was that "even if we avoid the tragedy of the commons, it will *only* be by recourse to the tragic necessity of Leviathan" (ibid., 220).

Empirical evidence suggests, however, that individuals facing social dilemmas in many cases develop credible *ex ante* commitments without relying on external authorities. Fishers, irrigators, herders, or other users appropriating from a CPR have repeatedly shown their capacity to organize themselves, to establish credible commitments, to monitor each others' behavior, and to impose sanctions on those who break their commitments. We will call all such users of CPRs "appropriators."[6]

Self-organized CPR institutions have been devised without reference to central authorities and sustained over long periods of time without enforcement by external agents.[7] Michael Hechter's

---

**[4]See for example, Taylor (1987), R. Harding (1971,1982), Campbell (1985), McLean (1981), Moore (1987).**

**[5]The term "social dilemma" refers to a broad class of games, including the PD game, where strategies leading to efficient joint outcomes are strictly dominated for each individual by strategies leading to deficient equilibria (see Dawes 1975, 1980). See Bianco and Bates (1990) for a theoretical analysis of the capabilities and limits of assigning leaders strong sanctioning powers, and Samuelson, et al. (1986) for an experimental investigation of the choice of a Leviathan-like mechanism to solve social dilemmas.**

**[6]See Gardner, Ostrom, and Walker (1990) for a detailed exposition of the terms we use and the models we have developed.**

**[7]See E. Ostrom (1990) for a summary of successful, failed, and fragile efforts to self-organize and govern, small-scale CPRs. See also, National Research Council (1986), Berkes (1989), etc., for summaries and synthesis of the extensive case literature.**

recent synthesis of self-organizing capabilities within many social groups provides evidence across other domains (Hechter 1987; Williamson 1975, 1985).[8] Experimental studies have repeatedly found that individuals placed in laboratory social dilemmas, who are allowed to communicate, consistently achieve better outcomes than are predicted by noncooperative game theory.[9]

Words alone, in some instances, enable individuals to arrive at precommitments with more credence than current game theory allows. While findings concerning the positive effect of communication in laboratory settings-covenanting without a sword—lead to optimistic predictions about the capacities of individuals to solve social dilemmas, many cases of self-organized arrangements in natural settings rely on sanctioning mechanisms (E. Ostrom 1990). In these instances, the "sword" is wielded by those involved rather than by external agents.

Past research has produced three anomalies from the standpoint of both Hobbesian theory and noncooperative game theory:

(1) In one-shot social dilemma experiments, communication alone leads to substantial improvements in joint outcomes.[10]

(2) In repeated social dilemma experiments, repeated communication alone leads to substantial improvements in joint outcomes.[11]

(3) In naturally-occurring, repeated social dilemmas, participants invest substantial time and effort monitoring and imposing sanctions on one another.[12]

---

**[8]Sec, also, Kreps (1990) for a useful synthesis of literature relevant to endogenous commitments within firm*.**

**[9]One-shot experimental social dilemmas without communication also frequently lead to higher levels of cooperation than predicted. See, for example, Dawes, et al. (1986).**

**[10]The extensive literature on one-shot communication in dilemma games includes: Bornstein and Rapoport (1988); Bornstein, et al. (1989); Caldwell (1976); Dawes, McTavish, and Shaklee (1977); Dawes, Orbell, and van de Kragt (1984); Dawes, van de Kragt, and Orbell (forthcoming); Edney and Harper (1978); Jerdee and Rosen (1974); Kramer and Brewer (1986); Braver and Wilson (1984,1986); Orbell, Dawes, and van de Kragt (1990); Orbell, van de Kragt, and Dawes (1988a,b); van de Kragt, et al. (1983, 1986).**

**[11]See Isaac and Walker (1988,1991) and E. Ostrom and Walker (1991) for a discussion of relevant literature.**

**[12]Acheson (1989), Alexander (1982), Berkes (1986, 1989), Davis (1984), McKean (1986), Netting (1981), E. Ostrom (1990), Pinkerton (1989), Schlager (1990), Tang (1991. forthcoming). Thomson (1977), Thomson, et al. (1986), Wade (1988).**

The results from one-shot and repeated social dilemmas with communication are consistent with one another. Communication improves outcomes. These findings are extremely robust. They have been confirmed by many researchers across differing experimental environments.

The results from field settings show that participants in social dilemmas do not rely on communication alone. They closely monitor each other and impose sanctions on those who do not conform to the rules they have devised. If communication alone were a fully reliable mechanism to overcome the gap between *ex ante* temptations and *ex post* promises, then one should not observe tune and effort being devoted to monitoring and sanctioning efforts.

Thus, prior research has generated some interesting puzzles. Given that social dilemmas lie at the foundation of the theory of the state and the theory of collective action, it is important to explore the independent and interactive effects of: (1) communicating (or, to use Hobbes's terms, covenanting), (2) sanctioning (or, the sword), and (3) communicating with options to sanction (covenants with a sword).

In this paper we undertake this exploration.[13] We manipulate treatments to examine: (1) communication alone (one-shot and repeated), (2) sanctioning alone, and (3) communication combined with the possibility of sanctioning. We construct a common constituent game which is the basis for all these manipulations. This game, an n-person CPR appropriation game is described and solved in the next section.

"The mathematical structure of all our experiments reported on in this paper is the same. The specific parameters vary somewhat among experiments, as is indicated in the text describing each aeries.

## II. Game-Theoretical Predictions

### A. The CPR Constituent Game

We will first specify the class of constituent CPR games from which we draw our designs. Assume a fixed number n of appropriators with access to the CPR. Each appropriates i has an endowment of resources e which can be invested in the CPR or invested in a safe, outside activity. The marginal payoff of the outside activity is normalized equal to w. The payoff to an individual appropriator from investing in the CPR depends on aggregate group investment in the CPR, and on the appropriator investment as a percentage of the aggregate. Let $x_i$ denote appropriator i's investment in the CPR, where $0 \le x_i \le e$. The group return to investment in the CPR is given by the production function $F(\sum x_i)$, where F is a concave function, with $F(0) = 0$, $F'(0) > w$, and $F'(ne) < 0$. Initially, investment in the CPR pays better than the opportunity cost of the foregone safe investment [$F'(0) > w$], but if the appropriators invest all resources in the CPR the outcome is counterproductive [$F'(ne) < 0$]. Thus, the yield from the CPR reaches a maximum net level when individuals invest some but not all of their endowments in the CPR.[14]

Let $x = (x_1,...,x_n)$ be a vector of individual appropriators' investments in the CPR. The payoff to an appropriator, $u_i(x)$, is given by:

$$u_i(x) = we \qquad \text{if } x_i = 0$$
$$w(e-x_i) + (x_i/\sum x_i)F(\sum x_i) \qquad \text{if } x_i > 0. \qquad (1)$$

(1) reflects the fact that if appropriators invest all their endowments in the outside alternative, they get a sure payoff (we), whereas if they invest some of their endowments in the CPR, they get a sure

**[14]Investment in the CPR beyond the maximum net level is termed "rent dissipation" in the literature of resource economics. This is conceptually akin to, but not to be confused with, me term "rent seeking," which plays an important role in political economy and public choice. For the latter, see Tullock (1980).**

payoff $w(e-x_i)$ plus a payoff from the CPR, which depends on the total investment in that resource $F(\ x_j)$ multiplied by their share in the group investment $(x_i/\ x_j)$.[15]

Let the payoffs (1) be the payoff functions in a symmetric, noncooperative game. Since our experimental design is symmetric, there is a symmetric Nash equilibrium, with each player investing $x_i^*$ in the CPR, where:

$$-w + (1/n)F'(nx_i^*) + F(nx_i^*)(x_i^*(n-1)/n^2) = 0. \qquad (2)$$

At the symmetric Nash equilibrium, group investment in the CPR is greater than optimal, but not all yield from the CPR is wasted.

There are several ways to interpret an equilibrium condition such as (2). One is in terms of disequilibrium, namely that any behavior not satisfying (2) will not persist over time, but will disappear. A second interpretation is in terms of equilibrium, namely that once behavior satisfies (2) that behavior persists over time. Neither of these interpretations says anything about the dynamics of behavioral change. A third and much stronger dynamic interpretation (2) is in terms of evolutionary stability. If behavior is being selected for according to the replicator equations, then (2) characterizes the unique dynamic stable equilibrium of the associated dynamical system (Hofbauer and Sigmund 1988).[16] A final interpretation is as a limited access CPR (see, for example, Clark 1980; Comes and Sandier 1986; and Negri 1989).[17]

---

**[15]This specification actually has a number of other possible interpretations. For instance, if one defines $F(\ x_j)/\ x_i = y$, and defines y to be a public good, then one has the payoff functions for a voluntary contribution mechanism as in Isaac and Walker (1988). Alternatively, one can define y in the same expression to be an externality, in which case one has payoff functions for Plott's experiments on externalities in product markets (Plott 1983). For further details, see Ledyard (1990).**

**[16]The proof of this result is available upon request.**

**[17]Consistent Conjectural Variations Equilibria may provide a useful method for a detailed analysis of individual subject behavior in these experiments. In the limited access version of the noncooperative CPR decision problem, full dissipation is predicted by nonzero consistent conjectures. See Mason, et al. (1988) for a discussion of consistent conjectures equilibria for the CPR experiment. See Walker, Gardner, and Ostrom (1991) for a discussion of several alternative theories that could be used to provide a solution to the core, constituent game.**

Compare this deficient equilibrium to the optimal solution. Summing across individual payoffs $u_i(x)$ for all appropriates i, one has the group payoff function u(x),

$$u(x) = nwe - w \quad x_i + F( \quad x_i) \qquad (3)$$

which is to be maximized subject to the constraints $0 \quad x_i \quad ne$. Given the above productivity conditions on F, the group maximization problem has a unique solution characterized by the condition:

$$-w + F'( \quad x_i) = 0. \qquad (4)$$

According to (4), the marginal return from a CPR should equal the opportunity cost of the outside alternative for the last unit invested in the CPR. The group payoff from using the marginal revenue = marginal cost rule (4) represents the maximal yield that can be extracted from the resource in a single period.

It is worth noting that both the Nash equilibrium investment and the optimum group investment do not depend on the endowment parameter e, as long as e is sufficiently large. For Nash equilibrium this seems especially counterintuitive, since large values of e represent high potential pressure on the CPR. Strategically, one of the most problematic aspects of a CPR dilemma is high over appropriation facilitated by high endowments. Alternative representations of this game do not suffer from this defect. In particular, the Von Neumann-Morgenstern characteristic function of the game representing this game does reflect the importance of e, as do some solutions based on the characteristic function such as the core. For example, for e large enough, the constituent game viewed as a cooperative game is a pure bargaining game.

**B. Finite Repetition of a CPR Constituent Game**

Denote the constituent game by X and let X be played a finite number of times. Typically, the repeated game has many equilibria. One often requires that a strategy specify equilibria play on subgame, the requirement of subgame perfection. If the constituent game has a unique equilibrium, then the finitely repeated game has a unique subgame perfect equilibrium (Selten 1971). Thus, equation (2) characterizes a finite sequence of equilibrium outcomes.

This prediction, like all predictions made in this paper, is based on the assumption of a finite game of complete information. Our experimental procedures assure that subjects know the game is finite.[18] Although we do not have complete control over our subjects' understanding of their decision task, the information we make available fulfills the requirements for complete information. Our complete information design, which requires the use only of subjects experienced in the constituent games, has a unique subgame perfect equilibrium. To the extent that we are unsuccessful in imparting complete information to our subjects, uniqueness of equilibrium is jeopardized. In the event of incomplete information, there is a bewildering multiplicity of game equilibria from which to select (Kreps et al., 1982). However, subjects experienced in playing the constituent game and given the level of information provided should approximate the ideal of a fully informed player.

**C. Communication and the Constituent Game**

When the constituent game X has a unique equilibrium x*, mere repetition does not create new equilibrium outcomes, nor does communication. Let c denote a communication strategy, in the communication phase C, available to any player. As long as saying one thing and doing another has no payoff consequences, then any strategy of the form (c, x*) is an equilibrium of the one-shot game (C,X). Furthermore, any sequence of games involving one-shot communication (C,X,X,...,X) or

[18]During recruitment, subjects we told they will participate in a 1-2 hour decision-making experiment. Although the exact endpoint is not revealed, it is explicitly bounded above. Further, all subjects are experienced and have thus experienced the boundedness of an experiment that lasted between 10 and 30 rounds.

involving repeated communication (C,X,C,X,...,C,X) has the same equilibrium outcomes as the constituent game repeated without communication. From this game-theoretic standpoint, covenants have no force. However, there is abundant empirical evidence that communication *does* make a difference to behavior. In this respect, noncooperative game equilibrium theory fails to address a well-known empirical regularity. We return to this problem below.

D. Sanctioning and the Constituent Game

Our sanctioning institution is represented formally using the following construction. Let s be a matrix of 0's and l's, where $s_{ij} = 1$ means that player i has sanctioned player j, and $s_{ij} = 0$ means that i has not sanctioned j. Row i of the matrix s codes all of player i's sanctioning behavior. As before, let x be a vector of individual investments in the CPR and $u_i(x)$ be i's payoff function in the game without sanctioning. Player i's payoff function in the game with sanctioning, $u_i(x,s)$, is given by:

$$u_i(x,s) = u_i(x) - f1 \sum_j s_{ij} - f2 \sum_j s_{ji}. \qquad (5)$$

The parameters f1 and f2 represent the cost of fining and the cost of being fined, respectively. The sum $s_{ij}$ is the total number of fines j levied by player 1, costing him f1 each; the sum $s_{ji}$ is the total number of times player i is fined, costing him f2 each.

Adding this sanctioning mechanism to our constituent game X produces a game X,S with a unique subgame perfect equilibrium. In a one-shot game with a unique Nash equilibrium x*, any sanctioning activity is costly and cannot lead to higher payoffs. Thus, the equilibrium of the one-shot game with sanctioning is the pair (x*,S*) = (x*,0), i.e., the equilibrium sanctioning matrix is the 0-matrix. At equilibrium, no one sanctions.

Now suppose that the one-shot CPR game with sanctioning is to be repeated a finite number of times T. Since the one-shot game has a unique equilibrium, the finitely repeated game still has a unique subgame perfect equilibrium, of the form (6):

In every period, play (x*,0).

In the event of any deviation from prescribed play, resume playing (x*,0) after the deviation. (6)

This equilibrium follows from backward induction. At the last period T, no deviation is profitable. At the next to last period T-l, given that no deviations will occur in the last period, then no deviation is profitable, and so on. Thus, according to any theory of repeated games that predicts subgame perfect equilibrium play, no sanctioning will take place (except by mistake) and the outcome each period will be the same as the one-shot Nash equilibrium. This conclusion continues to hold even if players do not know exactly when the game will end, but they know that the game will end in a finite number of periods (T). Given subjects' experience in this laboratory environment and particularly their knowledge that experiments last at most two hours, this assumption is valid. Suppose the most that T could be is 30 periods. Then if period 30 is reached, they would play the unique equilibrium at tune T=30. By backward induction, in period 29 they play the unique equilibrium, and so on for all possible periods.

Besides the unique subgame perfect equilibrium there is also a large class of imperfect equilibria. Let $z_i < x_i^*$ be the same for all i. Consider the repeated game strategy (7):

In every period except T, play (z,0).

In the event of any deviation, play $(x_i = 25, s = I)$ for one period, then resume playing (z,0). (7)

If no deviation took place in period T-1, play (x*,0) in period T.

Notice first of all that by putting less pressure on the CPR, individual payoffs improve (at least until the optimum level of investment is reached). Thus, individual players have an incentive to play (7). We claim that (7) represents an imperfect equilibrium. To show this, it suffices to show that no deviation from prescribed play pays. Let F(ne) be a very large negative number. For f1 and f2 large enough, a player who deviated optimally for one period would lose some positive amount, depending on the level of $z_i$, but in the next period would lose (l/n)F(ne) + f1 + f2 due to punishment from overinvestment and sanctions, as in (7). This threat we call the *dire threat,* as it is the worst threat imaginable for one period in our design. Given such a threat, it does not pay to deviate, even for one period. Finally, if a punishment is not called for in the last period, the endgame equilibrium is played in that period. This shows that (7) is an equilibrium. The imperfection of (7) lies in the fact that the trigger punishment-dumping all tokens into the resource, everybody fining everybody-is too harsh to be credible at the end of the game. But if it is not credible at the end of the game, then it is not credible one period from the end of the game . . . and so on down the slippery slope of backward induction.

It may seem peculiar that s = I, the identity matrix, so that everyone sanctions himself/herself when a deviation takes place. This representation is given just in the interest of symmetry-the repeated game has plenty of imperfect asymmetric equilibria. Indeed, taking any permutation of the identity matrix-for instance, 1 sanctions 2, 2 sanctions 3, and so on-also supports the same observed behavior, in terms of number of fines levied in case of a deviation. Notice that if mistakes are what cause deviations, then an equilibrium like (7) will generate n fines every time a mistake takes place, which is considerably more than the 0 fines generated by the subgame perfect equilibrium (6).

There is a large set of equilibria along the lines of (7), involving variation of the length of punishment (1 or more periods), the base level of investments $z_i$, and the direness of the one-period threat (dump not quite all tokens in the CPR, levy fines with some probability). In particular, by varying f1 and f2, we hoped to allow the subjects to find equilibria of the family (7) which involve punishments of the form ($z_i$,I)—that is to say, reduced investment in the CPR, but sanctions for everyone if a deviation occurs (see Jankowski 1990).

**E. Communication, the Constituent Game, and Sanctioning**

Our most complex design C,X,S consists of communication before and sanctioning after the constituent game. The payoff functions for this design are still given by (5) since communication per se has no payoff consequences and sanctioning does. Thus, we have the same subgame perfect equilibrium outcomes as for the design **X,S** (one-shot or finitely repeated). In addition, as in **X,S,** there exist many imperfect equilibria yielding higher payoffs than the perfect equilibrium outcomes. The role of communication could be to identify and signal the intention to play one of these imperfect equilibria. In this way, subjects may approach and maintain very high efficiencies within the confines of noncooperative game theory. This signalling mechanism would be especially effective if the opportunity to communicate was available every period so that subjects could verbalize punishments and intentions.

## III. The Laboratory Decision Environment

### A. Design

In our experimental investigation we have operationalized this CPR environment with eight appropriates (n = 8) and quadratic production functions F( $x_i$), where:

$$F(x_i) = a\, x_i - b(x_i)^2$$

with F'(0) = a > w and F'(nw) = a - 2bnw < 0.

For this quadratic specification, one has from (4) that the group optimal investment satisfies $x_i$ = (a-w)/2b. The CPR yields 0% on net when investment is twice as large as optimal, $x_i$ = (a-w)/b. Finally, from (2), the symmetric Nash equilibrium group investment is given by:

$Ex_i$ = ((n-l)/n)(a-w)/b.

This level of investment is between maximal net yield and zero net yield, approaching the latter as n gets large. One additional constraint that arises in a laboratory setting is mat the $x_i$ be integer-valued. This is accomplished by choosing the parameters a, b, d, and w in such a way that the predictions associated with $x_i$ are all integer valued.

In particular, we use the parameters shown in Table 1. These parameters lead to the predictions portrayed in Figure 1. A group investment of 36 tokens yields the optimal level of investment. This symmetric game has a unique Nash equilibrium with each subject investing 8 tokens in Market 2. At the Nash equilibrium, subjects earn approximately 39.5% of maximum net yield from the CPR. Once again, note that the Nash equilibrium and optimal investment are not affected by the level of endowments.

**B. Subjects and the Experimental Setting**

The experiments reported in this paper used subjects drawn from the undergraduate population at Indiana University. Students were volunteers recruited from principles of economics classes. Prior to recruitment, potential volunteers were given a brief explanation in which they were told only that they would be making decisions in an "economic choice" environment and that the money they earned would be dependent upon their own investment decisions and those of the others in their experimental group. All experiments were conducted on the NOVANET computer system at IU. The computer facilitates the accounting procedures involved in the experiment, enhances across experimental/subject control, and allows for minimal experimenter involvement.

# FIGURE 1

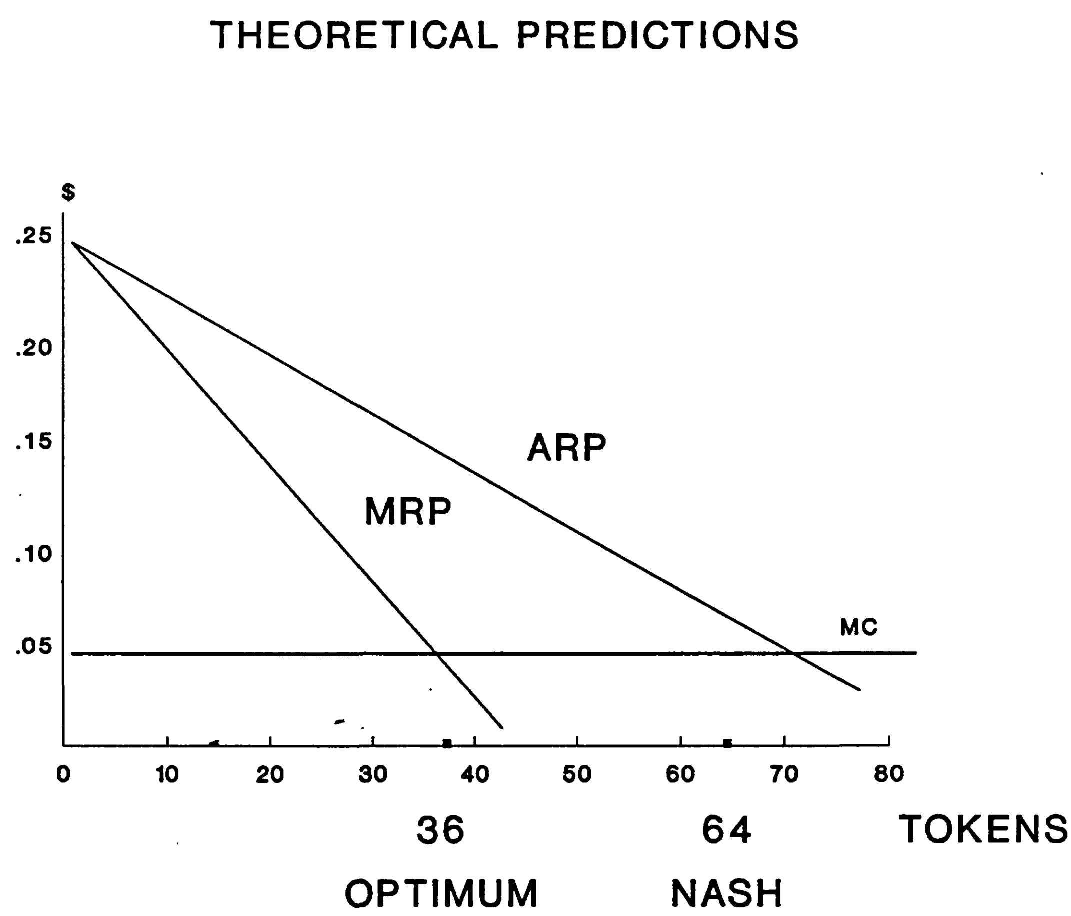

THEORETICAL PREDICTIONS
$
.25
.20
.15
.10
.05
ARP
MRP
MC
0
10
20
30
40
50
60
70
80
36
64
TOKENS
OPTIMUM
NASH

**Table 1**

**EXPERIMENTAL DESIGN BASELINE**
**Parameters for a Given Decision Period**

| Experiment Type: | LOW ENDOWMENT | HIGH ENDOWMENT |
|---|---|---|
| Number of Subjects | 8 | 8 |
| Individual Token Endowment | 10 | 25 |
| Production Function: Mkt.2* | $23(\ x_i)-.25(\ x_i)^2$ | $23(\ x_i)-.25(\ x_i)^2$ |
| Market 2 Return/unit of output | $.01 | $.01 |
| Market 1 Return/unit of output | $.05 | $.05 |
| Earnings/Subject at Group Max.** | $.91 | $.83 |
| Earnings/Subject at Nash Equil. | $.66 | $.70 |
| Earnings/Subject at Zero Net Yield | $.50 | $.63 |

* $Ex_i$ = the total number of tokens invested by the group in market 2. The production function shows the number of units of output produced in market 2 for each level of tokens invested in market 2.

** In the high endowment design, subjects were paid in cash one-half of their "computer" earnings. Amounts shown are potential cash payoffs.

## IV. The Baseline Game: No Covenants and No Swords

For comparison with the designs discussed later, the baseline experiments can be represented as an iterated series (of 20 rounds) of the constituent game X, shown in Figure 2. At the beginning of each experimental session, subjects were told that: (1) they would make a series of investment decisions, (2) all individual investment decisions were anonymous to the group, and (3) they would be paid their individual earnings (privately and in cash) at the end of the experiment. Subjects then proceeded at their own pace through a set of instructions that described the decisions.[19]

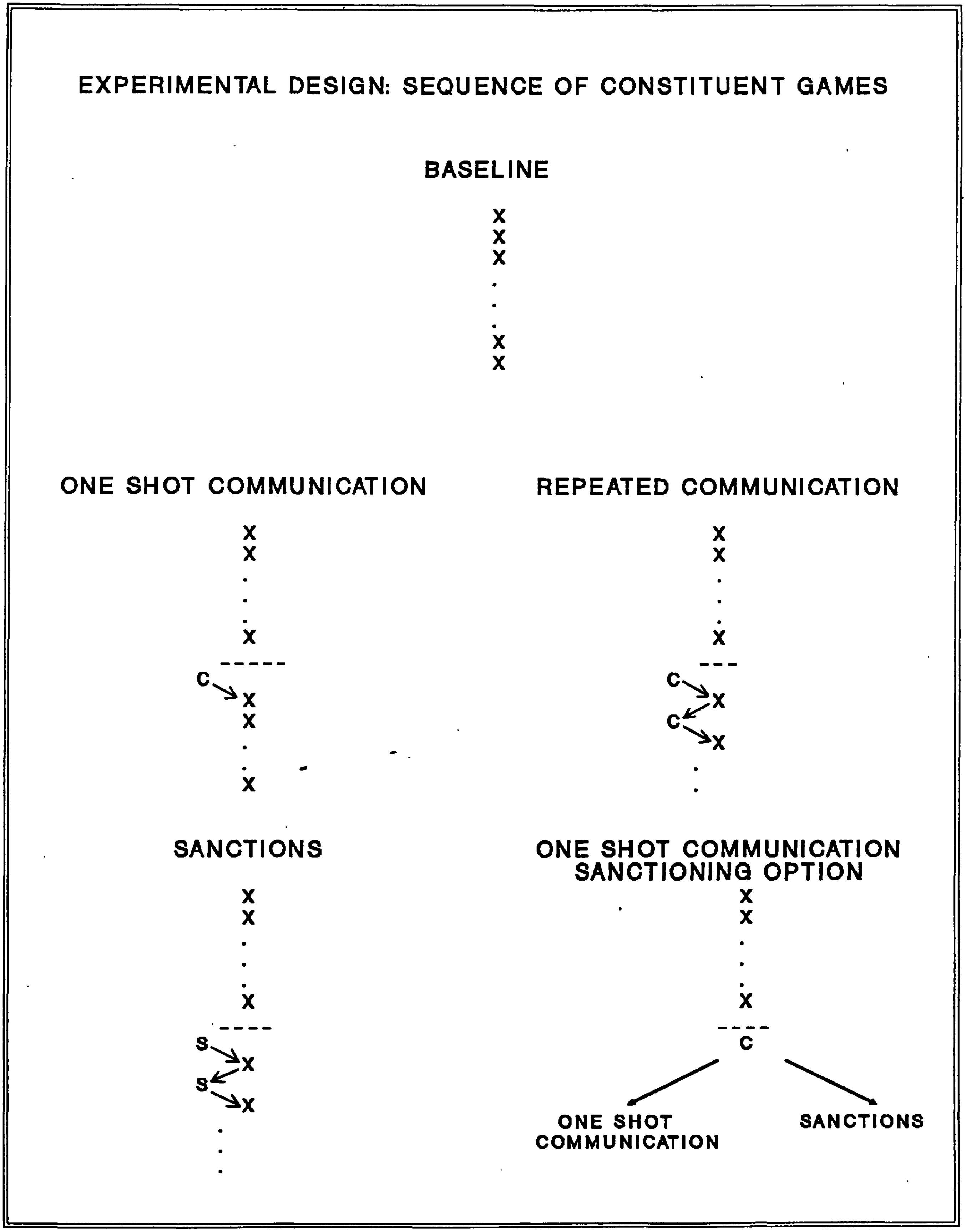
Subjects were instructed that in each period they would be endowed with a given number of tokens they could invest in two markets. Market 1 was described as an investment opportunity in which each token yielded a fixed (constant) rate of output and each unit of output yielded a fixed (constant) return. Market 2 (the CPR) was described as a market that yielded a rate of output per token dependent upon the total number of tokens invested by the entire group. The rate of output at each level of group investment was described in functional form as well as tabular form. Subjects were informed that they would receive a level of output from Market 2 that was equivalent to the percentage of total group tokens they invested. Further, subjects knew that each unit of output from Market 2 yielded a fixed (constant) rate of return. Figure 3 displays the actual information subjects saw as summary information in the experiment. Subjects knew with certainty the total number of decision makers in the group, total group tokens, and that endowments were identical. They did not know the exact number of investment decision periods. All subjects were *experienced,* i.e., had participated in at least one experiment using this form of decision environment.[20]

---

**[19]A complete set of instructions is available from the authors upon request. In all experiments reported in this paper, subjects were informed that their cash payoff would be one-half of the "lab" dollars earned in the experiment.**

**[20]The number of rounds in earlier experiments had varied from 10 to 20.**

# FIGURE 2


EXPERIMENTAL DESIGN: SEQUENCE OF CONSTITUENT GAMES
BASELINE
ONE SHOT COMMUNICATION
REPEATED COMMUNICATION
SANCTIONS
ONE SHOT COMMUNICATION
SANCTIONING OPTION
ONE SHOT
COMMUNICATION
SANCTIONS

# FIGURE 3

UNITS PRODUCED AND CASH RETURN FROM INVESTMENTS IN MARKET 2
commodity 2 value per unit = $ 0.01

| Tokens Invested by Group | Units of Commodity 2 Produced | Total Group Return | Average Return per Token | Additional Return per Token |
|---|---|---|---|---|
| 28 | 360 | $3.60 | $0.18 | $0.18 |
| 40 | 520 | $5.20 | $0.13 | $0.08 |
| 60 | 480 | $4.80 | $0.08 | $-0.02 |
| 80 | 240 | $ 2.40 | $0.03 | $-0.12 |
| 100 | -200 | $ -2.00 | $-0.02 | $-0.22 |
| 120 | -840 | $ -8.40 | $-0.07 | $-0.32 |
| 140 | -1680 | $-16.80 | $-0.12 | $-0.42 |
| 160 | -2720 | $-27.20 | $-0.17 | $-0.52 |
| 180 | -3960 | $-39.60 | $-0.22 | $-0.62 |
| 200 | -5400 | $-54.00 | $-0.27 | $-0.72 |

The table shown above displays information on investments in market 2 at various levels of total group investment. A similar table will be at your disposal during the experiment. Lets talk about the meaning of the information given in the table.

Press -NEXT- for the discussion
Press -BACK- to review

In the baseline experiments, subjects participated in a series of 20 decision periods. After each period, subjects were shown a display that recorded: (a) their profits in each market for that period, (b) total group investment in Market 2, and (c) a tally of their cumulative profits for the experiment. During the experiment, subjects could request, through the computer, this information for all previous periods. Players received no information regarding other subjects' *individual* investment decisions or concerning the number of iterations.[21]

A. Baseline Results

The baseline results are summarized in Table 2 and Figure 4. Table 2 displays information *regarding percentage of maximum net yield* actually earned by subject groups. The most striking observation comes from increasing token endowments from 10 to 25. Aggregating across all experimental periods, the average level of yields accrued in the low endowment (10 token) design equalled 37.2%. In contrast, the average level for the high endowment (25 token) design equalled -3.16%. The average tendencies for the first 20 decision rounds of the six experiments are presented in Figure 4.

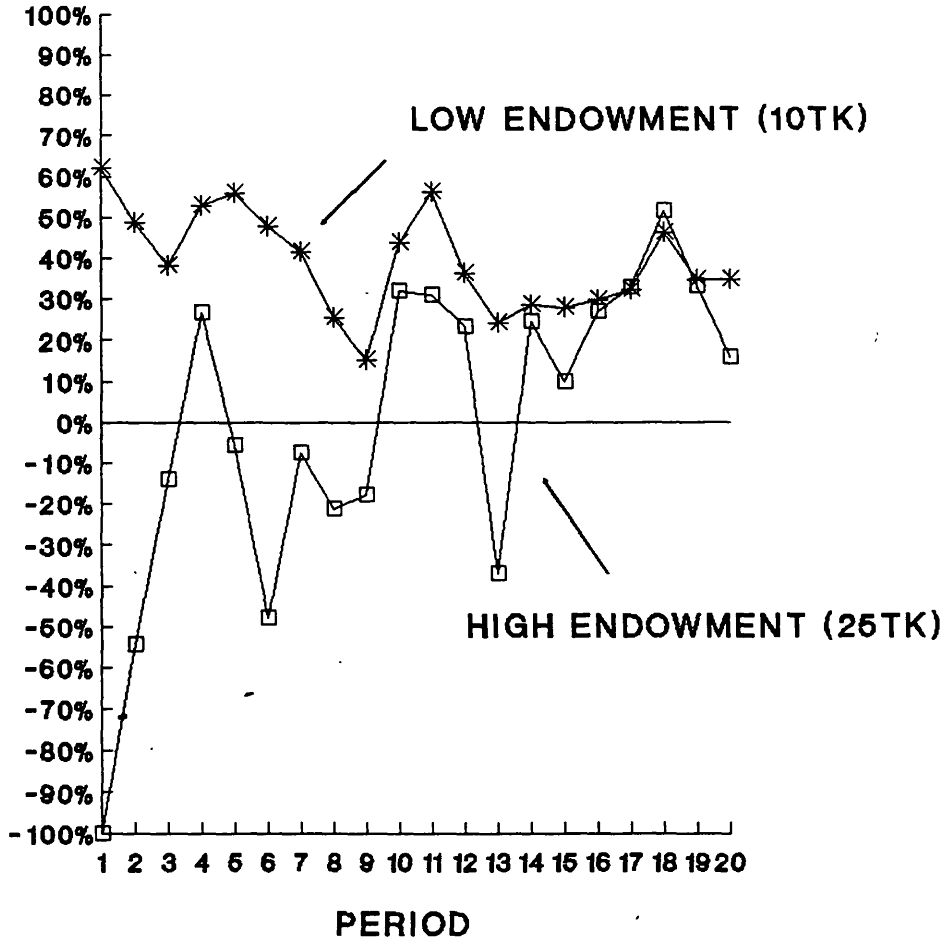
Several characteristics of the individual experiments are important. Investments in Market 2 are characterized by a "pulsing" pattern in which yield is reduced, at which tune investors tend to reduce their investments in Market 2 and yields increase. This pattern tends to recur across decision periods within an experiment. We did not find, however, symmetry across experiments in the amplitude or timing of peaks. For the high endowment experiments, the low points in the pulsing pattern were at yields far below zero. Over the course of the experiments, there was some tendency for the variance in yields to decrease. We saw no clear signs that the experiments were stabilizing. Further, we observed no experiments in which the pattern of individual investments in Market 2 stabilized at the

**"See Walker, Gardner, and Ostrom (1990) for details of this derivation. Maximum net yield = (Return from Market 2 minus the opportunity costs of tokens invested in Market 2)/(Return from Market 2 at MR=MC minus the opportunity costs of tokens invested in Market 2). Opportunity costs equal the potential return that could have been earned by investing the tokens in Market 1.**

# FIGURE 4

## THE EFFECT OF INCREASING INVESTMENT ENDOWMENT

YIELD AS A PERCENTAGE OF MAXIMUM
MEANS OF INDIVIDUAL EXPERIMENTS




NOTE: PERIOD 1 (26 TOKEN SERIES) • -162%

Table 2
Summary Results: Average Yield as a Percent of Maximum • Baseline Designs

ROUND

| Experimental Design | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 | 26+ |
|---|---|---|---|---|---|---|
| Baseline 10TK | 51.5 | 34.7 | 34.4 | 35.6 | 37.1 | 29.6 |
| Baseline 25TK | -42.5 | -12.4 | 10.3 | 32.0 | | |

Table 3
One-Shot Communication
Summary Results: Average Yield as a Percent of Maximum - Design **xx..cxx...x**

Round

| Exp.# | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 | 26+ |
|---|---|---|---|---|---|---|
| 1 | -47.5 | -20.2 | 88.9 | 88.9 | 85.1 | 83.0 |
| 2 | -73.2 | -15.6 | 45.3 | -0.08 | 11.8 | 31.7 |
| 3 | -02.0 | -02.4 | 88.0 | 48.1 | 30.5 | 61.0 |
| Mean | -40.9 | -12.7 | 74.1 | 45.4 | 42.5 | 58.6 |

Nash equilibrium. This is a behavioral result we have found consistently across over 100 experimental replications with alternative parameterizations. At the aggregate level, the Nash prediction best describes our data. However, at the individual level we have observed no case in which investments stabilize at the Nash prediction.

To sum up, as predicted by both Hobbesian theory and game theory, individuals acting independently in a CPR without communication or sanctioning do not solve the collective-action problem they confront.

## V. COVENANTS WITHOUT A SWORD-COMMUNICATION ALONE

Prior experimental evidence strongly supports the conclusion that communication in social dilemmas increases the frequency with which players choose and sustain joint income maximizing strategies, even when individual incentives conflict with such strategies. In this section we examine the results from several types of communication experiments.

### A. One-Shot Communication

Our first communication design was similar to that of the high endowment baseline game for the first ten repetitions of the constituent game. The only difference was that subjects received information on all individual decisions after each round. Information was given by subject number, thereby preserving anonymity. This added information had no impact on observed yields. At the end of the tenth round, the subjects were informed that they would have a single opportunity of ten minutes to discuss the decision problem. They were not to threaten one another or attempt side payments. [The instructions subjects read are reproduced in Appendix A.] After this ten-minute opportunity to communicate, the subjects returned to the constituent game which was then repeated up to 22 more times. The structure of the one-shot communication experiment is shown in Figure 2.

As discussed above, the equilibrium outcome for the one-shot communication game is for each individual to invest 8 tokens in the CPR, the same as in the baseline. The maximum yield is obtained if a total of 36 tokens are invested. Players were not allowed to invest fractional tokens and the symmetric strategy to obtain the maximum return is half-way between everyone investing 4 tokens and investing 5 tokens. Thus, discovering and agreeing upon a joint strategy is a cognitive, as well as a strategic, challenge in this environment. If the players were to decide to invest either 4 or 5 tokens each, they would obtain 99% of maximum net yield in either case.

The transcripts of the discussion during the communication round reveal that subjects perceived their problem as involving two tasks: (1) determining the maximal yield available and (2) agreeing upon a strategy to achieve that yield. The one-shot communication results are summarized in Table 3. As in Table 2, this table displays information *regarding percentage of maximum net yield* actually earned by subject groups.

The results of our three, one-shot communication experiments are mixed. In experiment 1, the group achieved over 82% of maximum net yield in all but 2 of 22 periods following communication. In experiment 2, communication had little efficiency-improving effects. Finally, in experiment 3, the group improved net yield significantly following communication, but could not sustain such behavior.

In experiment 1, the players agreed to invest 6 tokens each in the CPR. While this investment level is somewhat higher than optimal, the players still obtained 89% of the maximal return in rounds in which they complied with the agreement. The group complied perfectly until round 21, at which point compliance began to break down. In round 21, one subject invested 7 tokens. In round 22, three subjects invested 15, 7, and 7 tokens, respectively, producing a drop in yield to 59%. In round 23, one subject withdrew all tokens from the CPR, while the other 7 players returned to the agreed-upon 6 tokens. In all subsequent rounds, at least one player deviated from the agreement to invest 6 tokens. In round 28, the subject who had invested 15 tokens in round 22, invested 15 tokens again. Otherwise, all CPR investments ranged from 5 to 7 until termination in round 32.

In experiment 2, communication had no effect on yields. In the communication period, the subjects immediately identified an investment strategy of 5 tokens each. The subjects noted that one of them had invested 25 tokens in each of the first ten rounds. One subject surmised that this person could not be making too much, but little attention was paid to what they should do if this person persisted. Only one comment was made about their need to "stick to their agreement," and that comment was made by the 25-token investor (who remained anonymous throughout the experiment). In rounds 11 and 12, seven of the players invested the agreed upon 5 tokens, but the "heavy" investor from the first 10 rounds continued to invest 25 tokens. Thus, instead of earning 99% of maximal yield, the group earned only 56%. In round 13, one of these seven players doubled his investment in the CPR. This dropped the group yield to 35%. From round 14 to round 17, the group fluctuated between 20% and 55%. In round 18, several players increased CPR investment and yield plunged to -93%. When the experiment stopped after round 32, only two subjects were still investing 5 tokens in the CPR. No subject punished a defector by choosing to invest heavily in the CPR, as called for by trigger strategies employed in various folk theorems for infinitely repeated games. In fact, some subjects reduced their own investment levels in response to heavier investment by others. In 28 out of the 176 choices (or 16%), individuals invested less than the agreed-upon level of 5 tokens.

In experiment 3, communication had a positive but not a sustained impact on yields. The subjects wanted to adopt a strategy that would maximize yield, but had considerable difficulty identifying such a strategy. They finally decided upon a complex strategy to invest 3 tokens each in round 11 and one additional token each in rounds 12, 13, and 14. Depending on the information they obtained from these four rounds, their plan was that each player would continue to invest the number of tokens that had produced the highest return. Round 11 went according to plan. In round 12, seven subjects stuck with the plan, but one invested 21 tokens. In round 13, six did follow the plan and in round 14 all players invested 6 as agreed upon. In round 15, two players reduced investments to 3 and the other 6 invested 6 tokens-achieving a 97% yield. From round 16 onward, at least one person invested more than 6 in each round and the percentage of maximal returns plummeted to as low as -49%.

What is obvious from these three experiments is that a single communication period enables participants to begin the process of adopting a joint strategy. However, the incapacity to communicate repeatedly limits the long-run durability of their agreements.

**B. Repeated Communication**

Our second design involves repeated communication in both the low and high endowment settings. At the outset, the constituent game was repeated for 10 periods. After round 10, the players read an announcement (reproduced as Appendix B), informing them they will have an opportunity for discussion after *each* subsequent round. The players left their terminals and sat facing one another.[22] The structure of the repeated communication experiment is shown in Figure 2.

Summary data from the low endowment series is reported in Table 4.[23] These repeated communication experiments provide strong evidence for the power of face-to-face communication. Players successfully use the opportunity to: (a) calculate coordinated yield-improving strategies, (b) devise verbal agreements to implement these strategies, and (c) deal with nonconforming players through verbal statements. When allowed to communicate repeatedly, subjects greatly enhanced their

[22]As in the one-shot communication getting, each person was identified with a badge that was unrelated to their player number. This facilitated player identification in transcripts. If unanimous, players could forego discussion.
[23]These low endowment communication experiments were conducted very early in our research and used a modified 10 token payoff function

[23]These low endowment communication experiments were conducted very early in our research and used a modified 10 token payoff function for market 2 $(6 x_i) - .15(x_i)^2$. Yield as a percentage of maximum from experiments without communication using this payoff function closely parallel the yields observed in our 10 token low endowment baseline design. Across 20 decision periods, the difference in mean yields between experiments using these two alternative payoff functions for market 2 was only 6.4%, slightly higher in the low endowment baseline design presented in the text.

**Table 4**
**Repeated Communication - 10 Token Design**
**Summary Results: Average Yield as Percentage of Maximum -** XXX..CXCX...GX

Round

| Exp. # | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 |
|---|---|---|---|---|---|
| 1 | 25.6 | 25.7 | 96.1 | 100.0 | 100.0 |
| 2 | 34.7 | 21.0 | 100.0 | 97.2 | 100.0 |
| 3 | 33.2 | 23.5 | 98.6 | 98.7 | |
| 4 | 37.0 | 39.0 | 94.1 | 97.6 | 100.0 |
| Means | 32.6 | 27.3 | 97.2 | 98.4 | 100.0 |

**Table 5**
**Repeated Communication - 25 Token Design**
**Summary Results: Yield as a Percent of Maximum - Design** XXX..XCKCXCX

Round

| Exp.# | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 |
|---|---|---|---|---|---|
| 1 | 34.5 | -43.3 | 76.1 | 74.7 | 53.6 |
| 2 | 59.6 | 08.4 | 84.9 | 82.4 | 84.9 |
| 3 | 03.6 | -08.3 | 61.3 | 67.8 | 68.1 |
| Means | 32.6 | -14.4 | 74.1 | 75.0 | 68.9 |

joint yield and sustained this enhancement. In the low endowment environment, we identified only 19 defections from agreements out of 368 total decisions (a *5%* defection rate).[24]

The high endowment CPR game is a more challenging decision environment than the low endowment environment. While the equilibrium of the two games is identical, the disequilibrium implications of the 25-token game change considerably. With 25 tokens, as few as three subjects investing all of their tokens can essentially ruin the CPR (bring returns below w), while with 10 tokens it takes seven out of eight subjects to accomplish this much damage. In this sense, the 25-token environment is much *more fragile* than the 10-token environment. We were interested in exploring whether subjects could cope with this more delicate situation through communication alone.

Table 5 summarizes the data for the 25-token repeated communication experiments. In all three experiments joint yield increased dramatically. In experiment 1, however, the fragile nature of nonbinding agreements in this high endowment environment became apparent, particularly near the end of the experiment.

In experiment 1 the subjects disagreed about the best strategy—some arguing for investing 7 or 8 in the CPR and others arguing for 4 or 5. As the end of their first discussion period was announced, they rushed into a rapid agreement to "try 6 each and see what happens." All but one person kept to the agreement, with two extra tokens invested. After further discussion of whether 6 was the right amount, they again agreed upon this level of investment. One player ended the discussion by saying "Let's not get greedy. We just got to start trusting."

Fourteen extra tokens were invested in round 12, which produced a drop in their yield from 85% to 48% of maximum. When they next communicated, Player B announced:

---

**[24]We also conducted a series of costly communication experiments where subjects had to first pay for opportunity to communicate. The provision problem subjects faced in the costly communication experiments was not trivial and did in fact create a barrier. In all three experiments, the problem of providing the institution for communication diminished the success of either: (a) having the ability to develop a coordinated strategy and/or (b) dealing with players who cheated on a previous agreement. On the other hand, all groups succeeded to some degree in providing the communication mechanism and in dealing with the CPR dilemma. On average, net yield increased from 42% to 80%. See E. Ostrom and Walker (1991) for a detailed discussion.**

> This should be our last meeting-if we can't get some trust, we might as well go back and screw each other over. We could all make more money if we could stick together, but if some are going to do the others in, then, we just should go.

Rounds 13, 14, and 15 were close to the agreed-upon levels and yields were above 80% of maximum for each of these rounds. After round 15, the discussion period started off with:

| | |
|---|---|
| Player H | Not everyone is investing 6. |
| Player B | Evidently not. |
| Player C | Unless everyone keeps to it, it starts to get away from us. |
| Player H | Lets say we invest 6 again. Obviously, somebody is cheating, but what can we do? But the rest of us can just continue to invest 6. |

The players refused to talk after round 16. After yield dropped to 56% of maximum in round 17, the discussion started off with:

| | |
|---|---|
| Player E | Someone is getting a free ride so I say that we should just dump whatever we want into 2.[25] |
| Player H | But we screw ourselves too. |
| Player B | I think we should just turn it loose. |
| Player H | I am happy with continuing to invest 6. Yeah, someone is cheating, but that is the best we can do. Is it worth a dime or five cents? |
| Player E | [Obviously upset, shakes head and does not say anymore] |

The group in this experiment never again had perfect compliance.[26] But the threats to dump all the. tokens into the CPR were not carried out either. For five rounds, yields wavered between 72% and 80% of maximum. On the other hand, on the (unannounced) final round 23, their net yield plummeted to 11 % of maximum.

In experiment 2, the subjects mistook the optimal strategy. They adopted a group strategy of investing SO tokens in the CPR (2 subjects invest 7, 6 invest 6). They devised a complex rotation scheme and kept to it with only one exception throughout rounds 11 to 23.

---

**[25]The player is referring to market 2 (the CPR).**

**[26]Players H and E had followed the agreement through round 17; Player B had followed through round 16, but was one of the four individuals who invested more than the agreed-upon level in round 17. Player E invested 8 in round 18, but then returned to follow the agreement throughout the remaining rounds. Player H never deviated from the agreement. Player B alternated between 6 and 7 tokens in Market 2 after this discussion.**

When one subject invested 11 rather than the agreed-upon 6 tokens, no one knew who the errant person was or whether the additional investment came from a single player.[27] (Subjects had information only about total investments, as in the 10-token environment.) In the communication round following this defection, the dialogue went like this:

| | |
|---|---|
| Player A | Who did it? |
| Player C | Did someone get a little greedy? |
| Player E | We ended up with more tokens in Market 2. |
| Player C | But still the person who did put in the extra, they would not have made anymore, would they have? |
| Player E | Just a few darn cents above the rest of us. |
| Player A | Lets go back and try it the old way. |

After further discussion, Player A urged that "We should be able to keep this going a little longer," and Player F wondered whether the person who put in the extra tokens was "greedy or was it just an error." The subjects then returned to their terminals for three more rounds of perfect compliance with their rotation agreement.

In experiment 3, the participants again mistook the optimum strategy. They initially adopted a group strategy of investing 50 tokens (2 subjects investing 25 each, 6 investing 0) together with a rotation scheme. Several subjects articulated concerns about whether the experiment would continue long enough for them to complete the rotation, but they did agree on the system. When one subject put in 25 tokens for two rounds in a row, the information that 75 tokens had been invested in the CPR went without comment for one round. Once the rotation had been completed, the subjects discussed what to do now and whether the extra 25 tokens had been placed in error. They adopted a symmetric strategy of all investing 7 tokens in the CPR (20 tokens greater than optimum) which they held to with two small exceptions. When discussing these defections, one player asked "Why mess it up?" The implication was that small defections could lead to a worse outcome for all if they

**[27]This "defection" occurred in round 20. Since the baseline experiment was 20 periods, defections were more likely on the 20th round. That some defections come at this point points to the bounded nature of the experimental setting. That more defections do not come at this point or soon thereafter in the communication experiments is hard to explain using backward induction.**

continued. The implied threats worked relatively well in sustaining this suboptimal but yield-improving strategy.

In these three repeated communication experiments, subjects were able to obtain consistently higher payoffs than in the one-shot communication design, particularly after round 15. As shown in Table 5, subjects in the repeated, 25-token communication experiments obtained an average yield of 75% of maximum in rounds 16 to 20 as contrasted to 45% in the one-shot communication design. In rounds 20 to 25, the yields were 69% and 43%, respectively.

Subjects in the repeated communication setting were also able to keep their defection rates lower than in the one-shot setting. In the one-shot design, players invested more tokens in the CPR than agreed upon in 133 out of 528 opportunities (a defection rate of 25%), while the defection rate was 13% (42 out of 312) with repeated communication. Repeated communication enabled subjects to discuss defections and to cut the defection rate in half. In all communication experiments, subjects offered and extracted promises of cooperation, thereby increasing their joint yield significantly above that obtained prior to communication. Only in repeated communication did subjects develop verbal sanctioning mechanisms that enabled them to sustain consistently higher yields.

Communication discussions went well beyond discovering what investments would generate maximum yields. A striking aspect of the discussion rounds was how rapidly subjects, who had not had an opportunity to establish a well-defined community with strong internal norms, were able to devise their own agreements and verbal punishments for those who broke those agreements. These verbal sanctions had to be directed at unknown defectors, since subjects' decisions were anonymous. Subjects detected defection solely through aggregate investments. In many cases, statements like "some scumbucket is investing more than we agreed upon" were a sufficient reproach to change defectors' behavior. However, verbal sanctions were less effective in the 25-token environment.

These results are similar to those obtained in previous research in different but broadly similar environments.[28]

The evidence from the high endowment experiments explains why individuals in some natural settings do not rely on face-to-face communication alone. When the actions of one or a few individuals can be a strong disequilibrating force, individuals who have the capacity to agree to sanction one another as well as communicate with one another might well want to add the sword to a covenant. While the theoretical predictions are that individuals in such settings would not sanction one another, endogenous sanctioning is frequently observed in field settings. The question to which we now turn is whether sanctioning behavior will occur in a laboratory setting.

## VI. SWORD WITHOUT COVENANTS-SANCTIONING ALONE

This section summarizes a set of experiments focusing on sanctioning alone.[29] Experiments began like high endowment baseline experiments with the exception that after each round, subjects received *individual* data of all decisions. This information was given by subject number, thus maintaining anonymity. Our sanctioning mechanism required that each subject incur a cost (a fee) in order to sanction another. Subjects had no opportunity to discuss the sanctioning mechanism prior to its implementation. This created an experimental setting as close as possible to the noncooperative assumptions of no communication and no capacity to engage in enforceable agreements. In our first sanctioning design, after round 10, subjects were given an announcement summarized below. (For the full text see Appendix C.)

---

**[28]Isaac and Walker (1988,1990) found similar results for costless and costly communication in a public-good environment with symmetric payoffs. Similar to the results discussed above for costly communication, they found in their costly communication environment the success of face-to-face communication somewhat, but that, even with this reduction, the institution remained a successful mechanism for improving market efficiency.**

**[29]Earlier experiments focusing on sanctioning mechanisms without communication include Yamagishi (1986,1988).**

> Subjects were informed that in all remaining periods they would be given the opportunity to place a fine. Other subjects would have the same opportunity to fine them. Each subject could levy one fine at a specified fee. The player fined would pay a fine which was double the fining fee. It was possible for a single subject to be charged multiple fines. After each round, each subject filled out a fining form. These forms were collected and tallied by the experimenter, who then reported the results to each subject. Note that any subject who was fined did *not* know the identity of those who imposed the fine. At the end of the experiment, the experimenters subtracted from players' total profits the total of all fees and all fines.

The actual fees and fines that were used are reported in Table 6. After subjects read the announcement, questions regarding the implementation of the procedure were answered. No discussion was held on why the subjects might want to use the procedure or its possible consequences. Subjects did not know the end period of the experiment.

The principle results from our sanctioning experiments are summarized in Table 7.[30] Across all eight experiments, net yield rose from -38% before the imposition of sanctioning to 36% after. When one deducts the cost of operating the sanctioning institution, the overall improvement in subjects' payoffs is only 11%. Besides these quantitative results, we draw the following qualitative conclusions:

(1) significantly more sanctioning occurs than predicted by subgame perfection, and the frequency is inversely related to cost;

(2) sanctioning is primarily focused on heavy CPR investors;

(3) there is a nontrivial amount of sanctioning that can be classified as error, lagged punishment, or "blind" revenge.

We observed 176 instances of sanctioning across the eight experiments. In no experiment did we observe fewer than 10 instances. The frequency of sanctions is inversely related to the cost of

---

**[30]A comparison of the initial 10 rounds to the initial 10 rounds of baseline, suggests that the addition of anonymous information about individual decisions had no impact on investments.**

Table **6**
**SANCTIONING FEE AND FINE PARAMETERS**

| LEVEL OF FEES AND FINES | NUMBER OF EXPERIMENTS |
|---|---|
| $.057 $.10 | 2 |
| $.05/$.20 | 1 |
| $.40/$.80 | 2 |
| $.20/$.80 | 3 |

Table 7
Repeated Sanctioning
Summary Results: Average Yield as a Percent of Maximum - Design **xx..xsxs...xsx**

Round

| Exp.# | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 |
|---|---|---|---|---|---|
| 1 | -07.7 | -75.6 | 16.7 | 01.4 | 01.6 |
| 2 | -16.4 | -65.1 | 34.1 | 36.1 | -03.9 |
| 3 | -21.9 | -08.5 | 27.5 | 33.7 | 22.3 |
| 4 | -56.8 | -62.5 | 87.3 | 99.0 | 99.2 |
| 5 | -71.0 | 23.4 | 26.3 | 11.5 | 0.50 |
| 6 | -03.0 | 50.0 | 52.0 | 50.8 | 30.6 |
| 7 | 0.30 | -94.2 | 30.3 | 27.3 | 28.6 |
| 8 | -108.5 | -84.1 | 46.5 | 50.6 | 50.6 |
| Mean | -35.7 | -39.6 | 40.1 | 38.8 | 28.7 |

imposing the fine and dramatically increases with the stiffness of the fine. Further, our results, although reminiscent of equation (7), do not strictly support the conclusion that players were playing an equilibrium of this form. Except for one experiment, where net yield was over 95%, yields were too low and sanctioning levels were too high to be consistent with imperfect equilibria.

The second and third results relate to the reasons for sanctioning. From post-experiment interviews and personal observations, we offer four explanations for the higher-than-predicted level of sanctioning.

(a) One period punishment - the person fined was the highest or one of the highest investors in the previous period;

(b) Lagged punishment - the person fined was one of the highest investors in the CPR in either the no sanctioning periods or in earlier rounds of the sanctioning periods;

(c) Blind revenge — the person fined was a low CPR investor and was fined by a person fined in a previous period;

(d) Error - no obvious explanation can be given for the action (trembling hand).

In summary, 77% of all sanctioning is aimed at investors who in the previous round were above-average investors in the CPR. An additional 7% were aimed at players who had been heavy investors in the CPR in earlier (but not the most recent) rounds. We classify 5% as blind revenge and the remaining 11% as errors.[31]

The evidence from these experiments helps to explain why individuals in natural settings might not want to rely on sanctioning alone. Individuals who have the capacity to sanction one another without the ability to communicate with one another face an insuperable handicap to increasing their

---

**[31]A second set of sanctioning experiments was conducted as a check on the robustness of our original design. Readers of our earlier results conjectured that the lack of a significant improvement in rent accrual with the introduction of a sanctioning mechanism in our initial design could be due to a hysteresis effect tied to the decisions in the first 10 periods, periods in which there was no sanctioning mechanism. In our second design, three new experiments were conducted in which the sanctioning mechanism was introduced prior to the first decision period. In all three experiments the fee to fine ratio was $.40 to $.80. Subjects used fines repeatedly in all three experiments. On average, in each of the 20 period experiments, there were 17.3 fines placed per experiment. In summary, the results from our second set of experiments are very consistent with those from our first design. There was no persistent rent-improving behavior which can be tied to the introduction of the sanctioning mechanism. Examining efficiency when costs of fees and fines are incorporated leads to the same conclusion. In fact, it points to the composite result that fees and fines had a negative impact on net benefits accrued in the experiments. The mean level of efficiency in these three experiments was 67.8%. This contrasts to 75.9% in the three baseline experiments.**

payoffs. The final question to which we return is whether communication and sanctioning together foster sustainable high yields.

## VII. COVENANTS WITH A SWORD-COMMUNICATION AND SANCTIONING

Our last series of experiments investigates the consequences of allowing players a one-shot opportunity to communicate and decide to impose a sanctioning system upon themselves. Four of these experiments began like high endowment baseline experiments with the exception that after each round, subjects received *individual* data of all decisions. After round 10, subjects were given an announcement summarized below. For the full text see Appendix D.

> Subjects were informed that they would have a single 10-minute discussion period. At the end of 10 minutes they would vote on whether to institute a sanctioning mechanism and on the level of fines if they did institute one. The voting rule was strict majority with the status quo being no sanctioning.

Subjects in these experiments had participated in CPR experiments with and without the use of the sanctioning. Two additional high endowment experiments were conducted where the only difference was that the communication round occurred before the first constituent game.

The results from the six experiments conducted in this decision environment are summarized in Tables 8 and 9. Our discussion will include data on both yield and overall efficiency (payoffs as a percentage of maximum) to account for sanctioning costs. In summary, this environment leads to improvements in efficiency significantly above those we observed when sanctioning was imposed exogenously with no communication. The use and success of an opportunity to decide upon future strategies and sanctions, however, vary significantly over our six experiments.

In experiments 1 and 2, the subjects agreed to an investment strategy and a sanctioning mechanism that led to near optimal yield. Further, there was very little waste of resources in implementing the sanctioning mechanism. In experiment 1, no sanctions were used. Overall efficiency increased from an average of 70% in the first 10 periods to an average of 97% in the

**Table 8**
**One-Shot Communication with Sanctioning Option**

| Exp | Commu Round | Agree to use Fines | Fines Agreed Upon | Market 2: # of Tokens Agreed | Defec- tion Rate | Fining Rate |
|---|---|---|---|---|---|---|
| 1 | After 10th round | Yes | $0.10 | 6 | .00 | no need |
| 2 | After 10th round | Yes | $0.20 | 5 or 6 is best 8 too high | .04 | 1.0+ |
| 3 | After 10th Round | No | zero | no agreemt | na | na |
| 4 | After 10th Round | No | zero | 4 | .41 | na |
| 5 | First round | Yes | $1.00 | 5 | .00 | no need |
| 6 | First Round | Yes | $0.20 | rotate 5 and 6 | .04 | 1.0 |

Table 9
**One** Shot Communication with Sanctioning Option
Summary Results: Average Yield as a Percent of Maximum

Design **xx...xsxsx...sxsx or** XX...CXX...X

Round

| Exp.# | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 |
|---|---|---|---|---|---|
| 1 | -18.4 | -34.6 | 88.9 | 88.9 | 88.9 |
| 2 | -15.5 | 24.4 | 96.2 | 94.4 | 90.9 |
| 3 | 92.9 | 87.2 | 84.2 | 76.0 | 68.1 |
| 4 | -0.09 | -04.8 | 99.3 | 47.8 | -38.7 |
| Mean | 14.5 | 18.1 | 92.2 | 76.8 | 52.3 |

Design **cxsxsx....sxsx**

Round

| Exp.# | 1-5 | 6-10 | 11-15 | 16-20 |
|---|---|---|---|---|
| 1 | 98.8 | 98.9 | 98.8 | 98.8 |
| 2 | 94.8 | 95.1 | 94.6 | 81.9 |
| Mean | 96.8 | 97.0 | 96.7 | 90.4 |

periods following communication and implementation of the sanctioning mechanism. In experiment 2, there were several periods in which sanctions were employed. In nearly all cases, these fines were concentrated on players who, given their CPR investments, were deviating from the agreed-upon strategy. Overall efficiency increased from an average of 77% in the first 10 periods to an average of 96% afterwards.

The subject groups in experiments 3 and 4 treated the opportunity to communicate and devise a sanctioning mechanism very differently from the subjects in the other four experiments. Experiment 3 is an outlier. Out of more than a hundred experiments we have conducted, this is the only one where yields in the first 10 periods were essentially optimal. When given the opportunity to discuss the decision problem and choose a sanctioning mechanism, the group: (1) agreed that they did not need a mechanism and (2) agreed that no one should try to get "greedy"-i.e., invest too much in the CPR. The group held together for a few periods, after which yields began a gradual decline. This decline was due primarily to a gradual increase in CPR investments by two subjects. By round 25, yield had dropped to 56% of optimum and overall efficiency to 89% of optimum.

The first ten periods of experiment 4 exhibit the standard pulsing pattern in yields, averaging below 20%. When given the opportunity to communicate, most players in this group (not all) adamantly opposed the implementation of the sanctioning mechanism. The group discussed the optimal investment strategy; each subject agreed to invest 4 tokens in the CPR. The group members opposing the use of a sanctioning mechanism argued that: (1) it was too stressful; (2) fines couldn't be focused sufficiently, and at times "snowballed" into players fining other players with "revenge" in mind; and (3) a system of fines took money away from the group as a whole.

After the discussion, this group successfully followed their near optimal investment pattern for two rounds. In round 13, one subject increased his investment to 5 tokens. In round 14, this subject returned to 4 tokens but another subject invested 5 tokens. In round 14, both of these subjects marginally increased their CPR investments. The subsequent rounds showed a gradual increase in investments by virtually every player. By round 17 the group was back to a pattern of investments parallel to baseline conditions. Yields reached levels as low as -322% of maximum yield (overall efficiency as low as -3%). After the experiment ended, several of the subjects expressed the opinion that they should have established a sanctioning mechanism after all.

We have traced back to the specific sanctioning/no communication experiment in which each of these subjects participated. Of the 32 subjects in these four experiments, 18 voted for the implementation of a sanctioning mechanism. Of the 14 who voted no, 11 had participated in a prior sanctioning experiment in which the fee to fine ratio was \$.20/\$.80. Of the 18 who voted yes, only 3 had been in a \$.20/\$.80 design. We infer from this result that the high level of sanctioning activity in the \$.20/\$.80 design, the lack of overall efficiency gams, and the presence of blind revenge combined to impede the willingness of participants to choose a sanctioning mechanism. From this initial set of communication/sanctioning experiments, it is obvious that some subjects can find an optimal strategy, design a sanctioning mechanism, agree to impose it on themselves, and then achieve a high rate of conformance to their agreement without an external enforcer. On the other hand, prior negative experience with institutions that individuals view as punitive and inefficient is conducive neither to the design of better institutions nor to a willingness to agree to use them.

It is possible that the experience of the first ten rounds of the constituent game had an effect on mechanism choice. To examine this possible hysteresis effect, two additional experiments were conducted. In these two experiments, the opportunity to communicate and to adopt a sanctioning mechanism was available at the outset. The results of these two experiments (numbered 5 and 6) are found in Tables 8 and 9.

In both of these experiments, the subjects quickly agreed to an investment strategy and a sanctioning mechanism to punish defectors. In experiment 5, the subjects agreed to a strategy in which each invested 5 tokens in the CPR. The subjects in this experiment earned 99% of maximum yields in every round. The adopted sanctioning mechanism was never used.

The subjects in experiment 6 adopted a sanctioning mechanism with a $.10/$.20 fee to fine ratio. The subjects mistook 46 tokens for the optimal solution (a solution that still earned them 95% of maximum yields). More importantly, however, investing 46 tokens meant the group had to work out a rotation scheme in which two subjects invested 5 tokens and the remaining six subjects invested 6 tokens. This strategy was followed for two rounds. In round 3, one subject deviated by investing 6 tokens instead of 5, and was promptly fined by 2 others. The group returned to compliance until round 14. At this tune, the deviator from round 3 again deviated and was again fined twice. In round 19, a different subject deviated by investing 6 tokens. This subject was fined by one player. In this experiment, the subjects never explicitly agreed to a strategy that *all* players would fine a subject whose investment deviated from the agreed-upon strategy. In fact, most of the subjects did not impose a fine on those who deviated. It is worth noting, however, that the net benefit from investing 6 tokens instead of 5 is negative in the case in which there is at least one person fining.

The evidence from these experiments shows why individuals in field settings would want to have both communication and sanctioning available as instruments of self-governance. Communication alone may be sufficient, especially in environments where the actions of a few appropriates cannot imperil the stability of agreements. However, the evidence from our high endowment environment shows why individuals in field settings might not rely on communication alone. Monitoring and sanctions give agents a powerful set of institutions to avert destabilizing behavior in the social dilemma confronting them.

## VIII. CONCLUSIONS: REDISCOVERING THE POSSIBILITY OF SELF-GOVERNANCE

### A. The Anomalies Increase

Prior field and experimental research has produced three anomalies from the perspectives of classical Hobbesian theory and modern noncooperative game theory. From the field is the anomaly that:

1. Participants invest substantial tune and effort in self-monitoring and self-sanctioning in CPRs.

Consistent with the field, we find higher levels of sanctioning behavior by participants in a laboratory setting than predicted by noncooperative equilibrium. Also, our experimental findings are directly supportive of the other two anomalies reported in earlier laboratory-based studies of social dilemmas. Our findings (summarized in Table 10) show that:

2. Repeated CPR constituent games, with a one-shot communication opportunity, lead to a substantial improvement in joint outcomes.

3. Repeated CPR constituent games, with repeated opportunities to communicate, lead to a substantial improvement in joint outcomes.

Further, we find that:

4. Repeated communication opportunities lead to higher joint outcomes than one-shot communication.

Since the theoretical prediction for one-shot and repeated communication is the same as that for a finitely repeated constituent game without communication, repeating the opportunity for "mere jawboning" should make no difference. But it does.

For theorists who base predictions of increased levels of cooperation in repeated settings on the adoption of trigger strategies, our findings are not supportive:

5. In no experiment where one or more subjects deviated from an agreed-upon joint strategy, did the subjects then follow a trigger strategy of substantially increasing their investments in the CPR.

Table 10
Overall Results all Designs
Summary Results: Average Yield as a Percent of Maximum

ROUND

| Experimental Design | 1-5 | 6-10 | 11-15 | 16-20 | 21-25 | 26+ |
|---|---|---|---|---|---|---|
| Baseline 10TK | 51.5 | 34.7 | 34.4 | 35.6 | 37.1 | 29.6 |
| Baseline 25TK | -42.5 | -12.4 | 10.3 | 32.0 | | |
| Oneshot Communication 25TK | -40.9 | -12.7 | 74.1 | 45.4 | 42.5 | 58.6 |
| Repeated Communication 10TK | 32.6 | 27.3 | 97.2 | 98.4 | 100.0 | |
| Repeated Communication 25TK | 32.5 | -14.4 | 74.1 | 75.0 | 68.9 | |
| Sanction 25TK | -35.7 | -39.6 | 40.1 | 38.8 | 28.7 | |
| One Shot Communication 25TK No Sanction Chosen | 46.4 | 41.2 | 91.7 | 21.7 | 14.7 | |
| One Shot Communication 25TK Sanction Chosen[a] | -16.9 | -05.1 | 92.5 | 91.6 | 89.9 | 93.8 |
| One Shot Communication 25TK Sanction Chosen[b] | | | 96.8 | 97.0 | 96.7 | 90.4 |

[a] Communication and sanctioning choice occurred after round 10.
[b] Communication and sanctioning choice occurred after round 1; table displays this data beginning in round 11 for comparison purposes.

In fact, in some experiments where one or more subjects deviated from an agreed-upon joint strategy, some subjects subsequently *reduced* their investments in the CPR. When subjects discussed the problem of how to respond to one or more "free riders," they rejected the idea of dumping all of their tokens into the CPR.[32]

Three more anomalous findings were observed in the series of experiments discussed in this paper:

6. The predicted outcomes are the same in low and high endowment environments. However, the high endowment design exhibited greater instability, less efficiency, and less effective communication.

7. Contrary to predictions, subjects used the sanctioning mechanism even when they could not communicate.

Indeed, subjects directed most of their sanctions toward those who overinvested in the CPR.

8. According to subgame perfection, past experience should not affect the decision of whether to adopt a sanctioning mechanism. However, two of the six groups presented with the choice to adopt sanctions decided against. A high percentage of subjects in these two groups had experienced an environment in which a low-cost, punitive sanctioning mechanism was imposed and used extensively.

In summary, these experiments add to the growing body of evidence contrary to the predictions of classical Hobbesian thought and full information noncooperative game theory relying on subgame perfection. While noncooperative game theory has generated predictions that are confirmed in other settings, its predictions here are not. Social dilemmas, where individuals are able to communicate or create new institutions, require a reformulated theory.

## B. What are the Implications?

Two major implications follow from the results of this paper. The first relates to the world of policy. Policy makers responsible for the governance and management of small-scale, common-pool resources should *not* presume that the individuals involved are caught in an inexorable tragedy from

---

**[32]The dialogue reported on page 29 is typical of this type of discussion. The player resisting the idea of dumping all of the tokens into the CPR indicated that "we screw ourselves too."**

which there is no escape. Individuals may be able to arrive at joint strategies to manage these resources more efficiently. To accomplish this task they must have sufficient information to pose and solve the allocation problem they face. They must also have an arena where they can discuss joint strategies and perhaps implement monitoring and sanctioning. In other words, when individuals are given an opportunity to restructure their own situation they frequently, but not always, use this opportunity to make credible commitments and achieve higher joint outcomes without an external enforcer. We cannot replace the determinate prediction of "no cooperation" with a determinate prediction of "always cooperate." Our findings are consistent with the political theories of Hamilton, Hume, Madison, Montesquieu, Smith, and Tocqueville. The results challenge the Hobbesian conclusion that sovereigns must govern by being above subjects, by monitoring them, and by imposing sanctions on all who would otherwise not comply.[33]

The second major implication relates to the world of theory. In finitely repeated social dilemmas with a unique constituent equilibrium, a wide variety of treatments will not change the game equilibrium outcome. This theoretical robustness is deceiving, however, since such treatments as face-to-face communication, sanctioning, and enhanced resources lead to substantial behavioral differences. The theoretical assault on this problem requires a reformulation of the game model. Several avenues of research, including incomplete information and bounded rationality, show considerable promise. What is increasingly clear is that any approach must model how individuals actually conceive and solve the decision task they face.

[33]See V. Ostrom (1987,1989,1991) for an elucidation of an alternative theory to that of Hobbes.

## REFERENCES

Acheson, James M. 1989. "Where Have All the Exploiters Gone? Co-Management of the Maine Lobster Industry." In *Common Property Resources*, ed. Fikret Berkes. London: Belhaven Press, 199-217.

Alexander, Paul. 1982. *Sri Lankan Fisherman: Rural Capitalism and Peasant Society.* Australian National University Monographs on South Asia No. 7. Canberra: Australian National University.

Anderson, Jay M. 1977. "A Model of the Commons." In *Managing the Commons,* ed. Garrett Hardin and John Baden. San Francisco: Freeman, 38-41.

Aumann, R. J. 1978. "Survey on Repeated Games." In *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern,* ed. R. J. Aumann. Mannheim: Bibliographiches Institut.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Arnold, J.E.M., and J. G. Campbell. 1986. "Collective Management of Hill Forests in Nepal: The Community Forestry Development Project. In *Proceedings of the Conference on Common Property Resource Management,* National Research Council. Washington, D.C.: National Academy Press, 425-54.

Baines, Graham B.K. 1989. "Traditional Resource Management in the Melanesian South Pacific: A Development Dilemma." In *Common Property Resources,* ed. Fikret Berkes. London: Belhaven Press, 273-95.

Bell, Frederick W. 1972. "Technological Externalities and Common Property Resources: An Empirical Study of the U.S. Lobster Industry." *Journal of Political Economy* 80:148-58.

Bendor, Jonathan B., and Dilip Mookherjee. 1985. "Institutional Structure and the Logic of Ongoing Collective Action." Stanford, CA: Stanford University, Graduate School of Business.

Berkes, Fikret. 1986. "Marine Inshore Fishery Management in Turkey." In *Proceedings of the Conference on Common Property Resource Management,* National Research Council. Washington, D.C.: National Academy Press, 63-84.

Berkes, Fikret, ed. 1989. *Common Property Resources, Ecology and Community-Based Sustainable Development.* London: Belhaven Press.

Bianco, William T., and Robert H. Bates. 1990. "Cooperation by Design: Leadership, Structure, and Collective Dilemmas." *American Political Science Review* 84 (March): 133-47.

Bornstein, Gary, and Amnon Rapoport. 1988. "Intergroup Competition for the Provision of Step-Level Public Goods: Effects of Preplay Communication." *European Journal of Social Psychology* 18:125-42.

Bornstein, Gary, Amnon Rapoport, Lucia Kerpel, and Tani Katz. 1989. "Within- and Between-Group Communication in Intergroup Competition for Public Goods." *Journal of Experimental Social Psychology* 25:422-36.

Braver, Sanford L., and L. A. Wilson. 1984. "A Laboratory Study of Social Contracts as a Solution to Public Goods Problems: Surviving on the Lifeboat." Paper presented at Western Social Science Association, San Diego, April.

Braver, Sanford L., and L. A. Wilson. 1986. "Choices in Social Dilemmas. Effects of Communication within Subgroups." *Journal of Conflict Resolution* 30(1) (March):51-62.

Caldwell, Michael D. 1976. "Communication and Sex Effects in a Five-Person Prisoners' Dilemma Game." *Journal of Personality and Social Psychology* 33(3):273-80.

Campbell, R. 1985. "Background for the Uninitiated." In *Paradoxes of Rationality and Cooperation*, ed. R. Campbell and L. Sowden. Vancouver: University of British Columbia Press, 3-41.

Christy, Francis T., and Anthony D. Scott. 1965. *The Common Wealth in Ocean Fisheries.* Baltimore: Johns Hopkins Press.

Clark, Colin W. 1980. "Restricted Access to Common-Property Fishery Resources: A Game-Theoretic Analysis." In *Dynamic Optimization and Mathematical Economics*, ed P. T. Liu. New York: Plenum Press, 117-32.

Cordell, John C., and Margaret A. McKean. 1986. "Sea Tenure in Bahia, Brazil." In *Proceedings of the Conference on Common Property Resource Management*, National Research Council. Washington, D.C.: National Academy Press, 85-113.

Cornes, R., and Todd Sandier. 1986. *The Theory of Externalities, Public Goods, and Club Goods.* Cambridge: Cambridge University Press.

Cruz, Wilfrido. 1986. "Overfishing and Conflict in a Traditional Fishery: San Miguel Bay, Philippines." In *Proceedings of the Conference on Common Property Resource Management*, National Research Council. Washington, D.C.: National Academy Press, 115-35.

Dasgupta, P. S. 1982. *The Control of Resources.* Oxford: Blackwell.

Davis, Anthony. 1984. "Property Rights and Access Management in the Small Boat Fishery: A Case Study from Southwest Nova Scotia." In *Atlantic Fisheries and Coast Communities: Fisheries Decision-Making Case Studies*, ed. C. Lamson and A. J. Hanson. Halifax, Canada: Dalhousie Ocean Studies Programme.

Dawes, Robyn M. 1975. "Formal Models of Dilemmas in Social Decision Making." In *Human Judgement and Decision Processes: Formal and Mathematical Approaches,* ed. Martin F. Kaplan and Steven Schwartz. New York: Academic Press.

____________. 1980. "Social Dilemmas." *Annual Review of Psychology* 31:169-93.

Dawes, Robyn M., Jeanne McTavish, and Harriet Shaklee. 1977. "Behavior, Communication, and Assumptions about Other People's Behavior in a Commons Dilemma Situation." *Journal of Personality and Social Psychology* 35 (Jan.): 1-11.

Dawes, Robyn M., John M. Orbell, and Alphons J.C. van de Kragt. 1984. "Normative Constraint and Incentive Compatible Design." Eugene: University of Oregon, Department of Psychology.

Dawes, Robyn M., John M. Orbell, Randy T. Simmons, and Alphons J.C. van de Kragt. 1986. "Organizing Groups for Collective Action." *American Political Science Review* 80 (Dec.): 1171-185.

Dawes, Robyn M., Alphons J.C. van de Kragt, and John M. Orbell. forthcoming. "Not Me Or Thee but WE: The Importance of Group Identify in Eliciting Cooperation in Dilemma Situations: Experimental Manipulations." *Acta Psychologica.*

Edney, Julian J, and Christopher S. Harper. 1978. "The Commons Dilemma: A Review of Contributions from Psychology." *Environmental Management* 2(6):491-507.

Farrell, Joseph. 1987. "Cheap Talk, Coordination, and Entry." *Rand Journal of Economics* 18 (Spring):34-39.

Feeny, D. H. 1988. "Agricultural Expansion and Forest Depletion in Thailand, 1900-1975." In *World Deforestation in the Twentieth Century,* ed. John F. Richards and Richard P. Tucker. Durham, North Carolina: Duke University Press, 112-43.

Frohlich, Norman, and Joe A. Oppenheimer. 1970. "I Get By With a Little Help From My Fields." *World Politics* 23:104-20.

Frohlich, Norman, and Joe A. Oppenheimer. 1990. "Choosing Justice in Experimental Democracies with Production." *American Political Science Review* 84 (June):461-77.

Gadgil, Madhav, and Prema Iyer. 1989. "On the Diversification of Common-Property Resource Use by Indiana Society." In *Common Property Resources,* ed. Fikret Berkes. London: Belhaven Press, 240-72.

Gardner, Roy, Elinor Ostrom, and James Walker. 1990. "The Nature of Common-Pool Resource Problems." *Rationality and Society 2*:335-58.

Gordon, Scott. 1954. "The Economic Theory of a Common Property Resource: The Fishery." *Journal of Political Economy* 62:124-42.

Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science* 162:1243-248.

Hardin, Garrett, and John Baden. 1977. *Managing the Commons*. San Francisco: Freeman.

Hardin, Russell. 1971. "Collective Action as an Agreeable n-Presidents' Dilemma." *Behavioral Science* 16(5) (Sept.):472-81.

______________. 1982. *Collective Action*. Baltimore: Johns Hopkins University Press.

Harsanyi, John C., and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge: MIT Press.

Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley: University of California Press.

Hobbes, Thomas. 1960. *Leviathan*. Oxford: Basil Blackwell. Originally published in 1651.

Isaac, R. Mark, and James M. Walker. 1988. "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism." *Economic Inquiry* 24(4):585-608.

Isaac, R. Mark, and James M. Walker. 1991. "Costly Communication: An Experiment in a Nested Public Goods Problem." In *Contemporary Laboratory Research in Political Economy*, ed. Thomas Palfrey. Ann Arbor: University of Michigan Press, forthcoming.

Jankowski, Richard. 1990. "Punishment in Iterated Chicken and Prisoner's Dilemma Games." *Rationality and Society 2* (Oct.):449-70.

Jerdee, Thomas H., and Benson Rosen. 1974. "Effects of Opportunity to Communicate and Visibility of Individual Decisions on Behavior in the Common Interest." *Journal of Applied Psychology* 59(6):712-16.

Kaitala, Veijo. 1986. "Game Theories Models of Fisheries Management - A Survey." *In Dynamic Games and Applications in Economics*; ed. T. Basar. Berlin: Springer-Verlag, 252-66.

Kramer, R. M., and Marilyn M. Brewer. 1986. "Social Group Identity and the Emergence of Cooperation in Resource Conservation Dilemmas." In *Experimental Social Dilemmas*, ed. Henk A. Wilke, David M. Messick, and Christel G. Rutte. Frankfurt: Verlag Peter Lang, 205-34.

Kreps, David M. 1990. "Corporate Culture and Economic Theory." In *Perspectives on Positive Political Economy*, ed. James E. Alt and Kenneth A. Shepsle. New York: Cambridge University Press, 90-143.

Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. 1982. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* 27:245-52.

Ledyard, John O. 1990. "Is There a Problem with Public Good Provision." Pasadena: California Institute of Technology. Mimeo.

Lewis, Tracy R., and James Cowens. 1983. "Cooperation in the Commons: An Application of Repetitious Rivalry." Vancouver: University of British Columbia, Department of Economics.

Libecap, Gary D. 1990. *Contracting for Property Rights.* Cambridge: Cambridge University Press.

Levhari, S., and L. H. Mirman. 1980. "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution." *Bell Journal of Economics* 11:322-34.

Libecap, Gary D. 1989. "Distributional Issues in Contracting for Property Rights." *Journal of Institutional and Theoretical Economics* 145(1) (March):6-24.

Mason, C., Todd Sandier, and Richard Comes. 1988. "Expectations, The Commons, and Optimal Group Size. *Journal of Environmental Economics and Management* 15:99-110.

McCay, Bonnie M., and James M. Acheson. 1987. *The Question of the Commons. The Culture and Ecology of Communal Resources.* Tucson: The University of Arizona Press.

McHugh, J. L. 1972. "Jeffersonian Democracy and the Fisheries." In *World Fisheries Policy: Multidisciplinary Views,* ed. B. J. Rothschild. Seattle: University of Washington Press.

McKean, Margaret A. 1986. "Management of Traditional Common Lands (Iriaichi) in Japan." In *Proceedings of the Conference on Common Property Resource Management,* National Research Council. Washington, D.C.: National Academy Press.

McLean, Iain. 1981. "The Social Contract in Leviathan and the Prisoner's Dilemma Supergame." *Political Studies* 29(3):339-51.

Messerschmidt, Donald A. 1986. "Collective Management of Hill Forests in Nepal: The Community Forestry Development Project." In *Proceedings of the Conference on Common Property Resource Management,* National Research Council. Washington, D.C.: National Academy Press, 455-80.

Moore, Clement Henry. 1987. "Prisoners' Financial Dilemmas: A Consociational Future for Lebanon?" *American Political Science Review* 81(1) (March): 201-18.

Nash, John F. 1950. "The Bargaining Problem." *Econometrica* 18:155-62.

__________. 1951. "Non-Cooperative Games." *Annals of Mathematics* 54:286-95.

National Research Council. 1986. *Proceedings of the Conference on Common Property Resource Management.* Washington, D.C.: National Academy Press.

Negri, D. H. 1989. "The Common Property Aquifer as a Differential Game." *Water Resources Research* 25:9-15.

Netting, Robert McC. 1981. *Balancing on an Alp: Ecological Change and Continuity in a Swiss Mountain Community.* Cambridge: Cambridge University Press.

Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 85(6): 1356-375.

Ophuls, W. 1973. "Leviathan or Oblivion." In *Toward a Steady State Economy,* ed. H. E. Daly. San Francisco: Freeman, 215-30.

Orbell, John, Robyn Dawes, and Alphons van de Kragt. 1990. "The Limits of Multilateral Promising." *Ethics* 100 (April):616-27.

Orbell, John M., Alphons J.C. van de Kragt, and Robyn M. Dawes. 1988a. "Covenants Without the Sword." Paper prepared for the Conference on "Normative Underpinnings of Constitutional Order," Jackson Hole, Wyoming, September.

Orbell, John M., Alphons J.C. van de Kragt, and Robyn M. Dawes. 1988b. "Explaining Discussion-Induced Cooperation." *Journal of Personality and Social Psychology* 54(5): 811-19.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* New York: Cambridge University Press.

Ostrom, Elinor, and James Walker. 1991. "Communication in a Commons: Cooperation without External Enforcement." In *Contemporary Laboratory Research in Political Economy,* ed. Thomas Palfrey. Ann Arbor: University of Michigan Press, forthcoming.

Ostrom, Vincent. 1987. *The Political Theory of a Compound Republic: Designing the American Experiment.* 2nd. rev. ed. Lincoln: University of Nebraska Press.

______________. 1989. *The Intellectual Crisis in American Public Administration.* 2nd ed. Tuscaloosa: University of Alabama Press.

______________. 1991. *The Meaning of American Federalism: Constituting a Self-Governing Society.* San Francisco: ICS Press, forthcoming.

Panayoutou, T. 1982. *Management Concepts for Small-Scale Fisheries: Economic and Social Aspects.* Rome: FAO Fisheries Technical Paper No. 228.

Pinkerton, E., ed. 1989. *Co-operative Management of Local Fisheries. New Directions for Improved Management and Community Development.* Vancouver: University of British Columbia Press.

Plott, Charles R. 1983. "Externalities and Corrective Policies in Experimental Markets." *Economic Journal* 93:106-27.

Plott, Charles R., and Robert A. Meyer. 1975. "The Technology of Public Goods, Externalities, and the Exclusion Principle." In *Economic Analysis of Environmental Problems,* ed. Edwin S. Mills. New York: National Bureau of Economic Research.

Rapoport, Amnon. 1985. "Provision of Public Goods and the MCS Experimental Paradigm." *American Political Science Review* 79 (March): 148-55.

Rowe, Nicholas. 1989. *Rules and Institutions.* Ann Arbor: University of Michigan Press.

Runge, Carlisle Ford. 1984. "Institutions and the Free Rider: The Assurance Problem in Collective Action." *Journal of Politics* 46:154-81.

Samuelson, Charles D., David M. Messick, Henk A.M. Wilke, and Christel G. Rutte. 1986. "Individual Restraint and Structural Change as Solutions to Social Dilemmas." In *Experimental Social Dilemmas,* ed. Henk A.M. Wilke, David M. Messick, and Christel G. Rutte. Frankfurt: Verlag Peter Lang, 29-54.

Sandberg, Audun. 1990. "Fish for All on Solutions to CPR-Problems in North-Atlantic Environments." Working Paper. Bloomington: Indiana University, Workshop in Political Theory and Policy Analysis.

Schlager, Edella. 1990. "Model Specification and Policy Analysis: The Governance of Coastal Fisheries." Ph.D. dissertation. Bloomington: Indiana University, Department of Political Science.

Schotter, Andrew. 1980. *The Economic Theory of Social Institutions.* New York: Cambridge University Press.

Selten, Reinhard. 1971. "A Simple Model of Imperfect Competition Where 4 Are Few and 6 Are Many." *International Journal of Game Theory 2:*141-201.

Shubik, Martin. 1982. *Game Theory in the Social Sciences. Concepts and Solutions.* Cambridge, MA: MIT **Press.**

Sinn, Hans-Werner. 1984. "Common Property Resources, Storage Facilities, and Ownership Structures: A Cournot Model of the Oil Market." *Economica* 51 (Aug.):235-53.

Siy, Robert Y. 1982. *Community Resource Management: Lessons from the Zanjera.* Quezon City: University of the Philippines Press.

Sobel, Jordon H. 1985. "Utility Maximizers in Iterated Prisoner's Dilemmas." In *Paradoxes of Rationality and Cooperation,* ed. Richmond Campbell and Lanning Sowden. Vancouver: University of British Columbia Press, 306-19.

Tang, Shui Yan. 1991. "Institutional Arrangements and the Management of Common-Pool Resources." *Public Administration Review* 51(1) (Jan./Feb.):42-51.

Tang, Shui Yan. forthcoming. *Institutions and Collective Action: Governing Irrigation Systems.* San Francisco: ICS Press.

Taylor, Michael. 1987. *The Possibility of Cooperation.* Cambridge: Cambridge University Press.

Thomson, James T. 1977. "Ecological Deterioration: Local-Level Rule Making and Enforcement Problems in Niger." In *Decertification: Environmental Degradation in and Around Arid Lands,* ed. M. H. Glantz. Boulder: Westview Press, 57-79.

Thomson, James T., David Feeny, and Ronald J. Oakerson. 1986. "Institutional Dynamics: The Evolution and Dissolution of Common Property Resource Management." In *Proceedings of the Conference on Common Property Resource Management*, National Research Council. Washington, D.C.: National Academy Press, 391-424.

Tullock, Gordon. 1980. "Efficient Rent-Seeking." In *Towards a Theory of the Rent-Seeking Society,* ed. James M. Buchanan, Robert Tollison, and Gordon Tullock. College Station: Texas A&M Press.

van de Kragt, Alphons J.C., et al. 1983. "The Minimal Contributing Set as a Solution to Public Goods Problem." *American Political Science Review* 77 (March): 112-22.

van de Kragt, Alphons J.C., et al. 1986. "Doing Well and Doing Good as Ways of Resolving Social Dilemmas." In *Experimental Social Dilemmas,* ed. H. Wilke, D. Messick, and C. Rutte. Frankfurt am Main, FRG: Verlag Peter Lang, 177-204.

Wade, Robert. 1988. *Village Republics. Economics Conditions for Collective Action in South India.* Cambridge: Cambridge University Press.

Walker, James, Roy Gardner, and Elinor Ostrom. 1990. "Rent Dissipation in a Limited Access Common-Pool Resource: Experimental Evidence." *Journal of Environmental Economics and Management* 19:203-11.

Walker, James, Roy Gardner, and Elinor Ostrom. 1991. "Rent Dissipation and Nash Disequilibrium in Common Pool Resources: Experimental Evidence." In *Game Equilibrium Models,* ed. Reinhard Selten. Berlin: Springer-Verlag, forthcoming.

Williamson, O. E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

______________. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York: Free Press.

Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51(1): 110-16.

______________. 1988. "Seriousness of Social Dilemmas and the Provision of a Sanctioning System." *Social Psychology Quarterly* 51(l):32-42.

# APPENDIX A

**ANNOUNCEMENT**

Some participants in experiments like this have found it useful to have the opportunity to discuss the decision problem you face. You will be given ten minutes to hold such a discussion. You may discuss anything you wish during your ten-minute discussion period, with the following restrictions.

1) **You are not allowed to discuss side payments.**
2) **You are not allowed to make physical threats.**
3) **You are not allowed to see the private information on anyone's monitor.**

Since there are still some restrictions on communication with one another, we will monitor your discussions.

We will be tape recording your discussions for our records.

# APPENDIX B

**ANNOUNCEMENT**

Sometimes in previous experiments, participants have found it useful, when the opportunity arose, to communicate with one another. We are going to allow you this opportunity between periods. There will be some restrictions.

1) **You are not allowed to discuss side payments.**
2) **You are not allowed to make physical threats.**
3) **You are not allowed to see the private information on anyone's monitor.**

Since there are still some restrictions on communication with one another, we will monitor your discussions between periods. To make this easier, we will have all discussions at this site.

APPENDIX C

**ANNOUNCEMENT**

In all remaining periods you will be given the opportunity to place a "fine" on another player in the experiment and/or have other players place a "fine" on you.

If you are willing to pay a fee of $.05, you will be able to impose a fine on one of the other players. If you decide to fine a player, that player must pay a fine of $.10. Any player that is fined must pay a fine of $.10 for **each** fine. For example: Assume 3 players decide to fine player X. Player X would then have to pay a fine of 3 tunes $.10 or $.30.

**Here is how the procedure will work.**

**AFTER EACH ROUND:**

1) Each player will be asked to fill out a "FINE SHEET." See the example "FINE SHEET," which has been given to you. Note that on the fine sheet you are able to impose a fine on only one player following any given decision round. After all players have filled out a "FINE SHEET," the fine sheets will be collected privately by the experimenter.

2) The experimenters will tally all fines and report the results back to each player. Note that any player who is fined does not know the identity of the other player or players who imposed the fine. **At no time during or after the experiment will players know the actual identity of other players, nor will they know which players chose to impose fines or which players were fined by other players.**

3) Each player will be asked to fill out his/her "ACCOUNTING SHEET." The Accounting sheet is used by you to keep tabs on the number of times you chose to pay the $.05 fee to have another player fined and/or the number of times other players imposed a $.10 fine on you. The experimenters will also keep track of all $.05 fees and all $.10 fines for each player.

4) At the end of the experiment, the experimenters will subtract from your total profits the total of all $.05 fees and all $.10 fines that you must pay. That figure will then be multiplied by one half and the result paid privately in cash.

Example for Player X: Initial Profits = $18.00
Total Fees and Fines = $2.50
Final Profits = $15.50
Cash Paid = (1/2) x $15.50 = $7.75

Note: Since we are paying one half of "Final Profits," you can interpret the payoff as though you are receiving one half of your profits from investing in markets 1 and 2 minus one half of all $.05 fees and all $.10 fines.

**ARE THERE ANY QUESTIONS? YOUR FIRST OPPORTUNITY TO PLACE A FINE ON ANOTHER PLAYER OR HAVE A FINE/OR FINES PLACED ON YOU WELL OCCUR AFTER THE INVESTMENT DECISIONS ARE REPORTED FOR THE NEXT PERIOD (PERIOD 11).**

APPENDIX D

**ANNOUNCEMENT**

Some participants in experiments like this have found it useful to have the opportunity to place fines on players who make very large investments in Market 2. You will be given ten minutes to discuss whether you would like to do this or not. Specifically, you will need to agree to whether you use a fine system at all. If you agree to this, you will need to set the amount of the fine that each person can assess another player: 5 cents, 10 cents, 15 cents, 20 cents, or more. The above decision will need to be made in light of the rule that we will charge each person who wishes to fine someone else a fee of one half of the fine. So, if you were to decide on a fine of 10 cents, a player would have to pay 5 cents in order to impose that fine. That would mean that a participant who wanted to fine another player after a round would be given a slip like the one you now have at your terminal. The player would indicate a willingness to pay a fee of X so that an amount of 2X would be deducted from the total earnings of another specific player.

You may discuss anything you wish during your ten minute discussion period, with the following restrictions.

**1) You are not allowed to discuss side payments.**
**2) You are not allowed to make physical threats.**
**3) You are not allowed to see the private information on anyone's monitor.**

Since there are still some restrictions on communication with one another, we will monitor your discussions between periods. To make this easier, we will have all discussions at this site.

We will be tape recording your discussions for our records.

At the end of the discussion period, we will ask you to vote on the following:

1. What level of fines-5, 10, 15, 20 cents, etc.-you might want to use.

2. Whether you wish to use that fining system at all.

If at least 5 of you can agree to a level for the fine and at least 5 of you vote yes in regard to having that fining system, we will then operate that system for you. If less than 5 agree to a fining system, we will return to the experiment and continue it as it was run for the first 10 rounds.

You will have one and only one opportunity of 10 minutes to discuss this option.