

SANCTIONING BY PARTICIPANTS IN COLLECTIVE ACTION PROBLEMS

by

James Walker and Roy Gardner
Department of Economics

and

Workshop in Political Theory and Policy Analysis

Elinor Ostrom

Department of Political Science

and

Workshop in Political Theory and Policy Analysis

January 11, 1991

Financial support from the National Science Foundation (Grant SES-4843901) is gratefully acknowledged. All data are stored on permanent PLATO disk files and are available on request. Send inquiries to Professor James M. Walker, Department of Economics, Ballantine 901, Indiana University, Bloomington, Indiana, 47405.

A previous version was presented at the Conference on Experimental Research on the Provision of Public Goods and Common-Pool Resources, May 18-20, 1990, Indiana University, Bloomington and the 1990 Annual Meetings of the American Political Science Association, the San Francisco Hilton, August 30 through September 2, 1990. Copyright by the American Political Science Association. We would like to thank Sheryl Ball, Vincent Ostrom, Charles Plott, and participants in the experimental conference at Indiana University for their useful comments.

SANCTIONING BY PARTICIPANTS IN COLLECTIVE ACTION PROBLEMS

Why . . . should the individual member of a community go to the trouble of punishing a free rider when he could be a free rider on the sanctioning efforts of others? (Taylor, 1987: 22).

I INTRODUCTION

Current models of collective action based on noncooperative game theory do not explain the variation of behavior and outcomes observed in many collective action settings. We focus on a particular subset of these, where multiple appropriators withdraw resource units from a common-pool resource (CPR) (Gardner, Ostrom, and Walker, 1990).¹ Considerable field research has generated empirical findings from CPRs in diverse cultural and governmental settings. Evidence from these studies can be classified into at least four broad categories.

1. Clearly Pareto inferior outcomes -- appropriators' behavior has led to strictly Pareto inferior outcomes and, in some cases, to the destruction of the resource upon which their livelihood depends.²
2. Long-lived, endogenous monitoring and sanctioning systems — appropriators have designed rules regulating the entry and appropriation from a CPR which are enforced by the appropriators themselves.³ Outcomes may not reach the Pareto frontier, but must be close enough to justify continued investments in monitoring and sanctioning behavior.

¹ We rely on Plott and Meyer (1975) to refer to all individuals withdrawing resource units as appropriators.

² See Christy and Scott (1965); Bell (1972); McHugh (1972); Sandberg (1990).

³ See Netting (1981); Berkes (1986); Siy (1982); McKean (1986); Libecap (1990); Ostrom (1990).

3. Short-lived endogenous monitoring and sanctioning systems — appropriators cease to monitor and sanction after an exogenous shock, such as a major change in factor prices or a dramatic increase in population.⁴
4. Short-lived exogenous monitoring and sanctioning systems -- external authorities impose rules regulating entry and appropriation, but fail to enforce these rules.⁵

The first category of field results is not surprising. These results are consistent with predictions derived from widely accepted theories of collective action, particularly applied to natural resource settings. The findings in the other three categories are, on the other hand, all surprising, but for different reasons. Findings in the second category illustrate the feasibility of appropriators designing their own institutions and investing time and effort in monitoring and sanctioning behavior. This is not a typical result associated with subgame perfect equilibrium theory. Findings in the third category illustrate that endogenously designed systems can collapse. This is not so surprising given the catastrophic nature of the shocks involved in many of these settings. What is surprising from a theoretical perspective is that such endogenous institutions existed without external enforcers for substantial periods of time prior to collapse. Findings in the fourth category illustrate that the remedies so often prescribed for solving CPR problems are frequently ineffective.

This paper will focus on explaining monitoring and sanctioning, since these activities are crucial to an explanation of the findings in all four categories discussed above. In

⁴ See Alexander (1982); Cruz (1986); Baines (1989).

⁵ See Davis (1984); Feeny (1988); Thomson (1977); Thomson, Feeny, and Oakeron (1986); Arnold and Campbell (1986); Messerschmidt (1986); Gadgil & Iyer (1989); Cordell & McKean (1986); Cruz (1986); Dasgupta, 1982; Panayoutou (1982); Pinkerton (1989).

Section II we summarize two examples of field settings that fall into the second category to provide a more detailed view of what this behavior looks like in natural settings. In Sections III and IV we move from field settings into an experimental laboratory setting where a substantial level of control over relevant parameters is achieved. Section III provides a baseline situation of limited access CPRs where appropriators cannot monitor or sanction. In Section IV, we analyze experiments where appropriators monitor each others' behavior and sanction one another if they are willing to expend resources to do so. We find that subgame perfect equilibrium theory does not explain observed sanctioning behavior the field or experimental settings of limited access CPRs.

II. MONITORING AND SANCTIONING IN FIELD SETTINGS

Anthony Davis (1984) provides a good description of the rules that characterize many Nova Scotian inshore fishing grounds. In these settings, fishers have defined the fishing grounds that are assigned to each village and defend these boundaries tenaciously.

... a Port Lameron Harbor fisherman, after setting his longline gear, watched a fisherman from a neighboring harbor set his gear close to and, on occasion, across his line. Subsequently, the Port Lameron Harbor fisherman contacted the "transgressor" on the citizen band radio to complain about this behavior. Other Port Lameron Harbor fishermen who were "listenin' in" on the exchange demonstrated support for their compatriot by adding approving remarks once the original conversation had ended. The weight of this support, coupled with the implied threat of action, i.e., "cuttin' off" the offender's gear, compelled the erring fisherman to offer his apologies (Davis, 1984: 147).

Further down the eastern coast of North America, the Maine lobster fishers are also noted for their willingness to monitor and sanction those not following accepted practices. James

Acheson, who has studied Maine fishers over long periods of time indicates that incursions by outside fishermen are not universally sanctioned. "A high-status fisherman [from a neighboring village] with many potential allies might get away with his incursion for weeks or months, while another man might be sanctioned almost immediately" (Acheson, 1989: 203). But almost all incursions that last beyond a few days will be sanctioned.

Usually the interloper is first warned by tying two half-hitches around the spindle of his buoys, or conspicuously leaving the doors on his traps open. If he persists in fishing in the area of another gang, his traps will then be destroyed -- usually by cutting the buoy line so that the trap is irretrievably lost. Usually the man or men defending the boundary will not 'cut off' any more traps than is necessary to force the interloper to move. Touching another man's traps is considered somewhat shameful, and is certainly illegal. In many cases, the interloper will retreat without a fight, in the knowledge that he has violated a boundary. Sometimes the victim will retaliate and a small 'cut war' will occur, in which men destroy the traps of guilty and innocent alike. Large-scale lobster wars', involving dozens of fishermen and the destruction of hundreds of traps, do occur — but they are very rare, occurring perhaps only once in a decade.

Information about trap-cutting incidents rarely escapes from the harbor gangs involved. The police or wardens are rarely called (Acheson, 1989: 203).

Acheson describes an endogenous form of sanctioning that is actually illegal in the broader legal system. It is not unusual to find inshore fishermen devising rules to limit access to their fishery and developing modes of enforcement that conflict with the legal rules of the larger jurisdiction of which they are formally a part. This makes it more difficult to explain why fishers would take the time and effort involved to sanction others. In addition to the cost of the activity itself and the risk of physical retaliation, there is the additional risk of a penalty imposed by external authorities for sanctioning activity. Given the implicit belief in the efficacy of external enforcers, it is hard to explain why police or wardens are rarely called in to enforce the laws against trap cutting!

Understanding internally imposed monitoring and sanctioning behavior more thoroughly could be a key to understanding endogenous solutions to collective action dilemmas. Using controlled laboratory experiments, the next two sections examine more closely existing theories of CPR equilibrium in the absence of and presence of sanctioning and monitoring.

III. BASELINE EXPERIMENTS: Theory, Design, and Results

Equilibria in the Baseline CPR Game

In this section, we examine four benchmark theories which have been used for describing behavior in open and limited access CPR situations. In our design there are a fixed number n of appropriators with access to the CPR. Each appropriator i has an endowment of resources w 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 c . The payoff to an individual appropriator from investing in the CPR depends on how much group investment there is in the CPR, and how much an appropriator invests as a percentage of the group total. Let x_i denote appropriator i 's investment in the CPR, where $0 \leq x_i \leq w$. 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 $F(0) = 0$, $F'(0) > c$, and $F'(nw) < 0$. Initially, investment in the CPR pays better than the opportunity cost of the foregone safe investment, but if the appropriators invest their resources in the CPR beyond some optimal quantity, the outcome is counter productive to the group at the margin. In the terminology of resource economics this would be "rent dissipation," where rents are defined as the

return to resources employed in the CPR above their opportunity costs.⁶

Let $x = (x_1, \dots, x_n)$ be a vector of individual appropriators' investments in the CPR.

The payoff to an appropriator, $u_i(x)$, is given by :

$$\begin{aligned} u_i(x) &= cw && \text{if } x_i = 0 \\ c(w-x_i) + (x_i/\sum x_j)F(\sum x_j) && \text{if } x_i > 0. \end{aligned} \quad (1)$$

(1) reflects the fact that if an appropriator invests all his resources in the outside alternative, he gets a sure payoff cw , whereas if he invests some of his resources in the CPR, he gets a sure payoff $c(w-x_i)$ plus a payoff from the CPR which depends on the total investment in that resource $F(\sum x_j)$ multiplied by his share in the group investment $(x_i/\sum x_j)$.⁷

The first theory proposed for organizing behavior in a CPR, which we denote T1, asserts that the appropriators overcome their first-order collective action problem and form an appropriator group to exploit the CPR optimally. Such an appropriator group will maximize the joint payoff from the resource. Summing across individual payoffs $u_i(x)$ for all appropriators i , one has the group payoff function $u(x)$,

$$u(x) = ncw - c\sum x_i + F(\sum x_j) \quad (2)$$

⁶ We use the term "rent dissipation" in the sense it is used in resource economics, where rents from a natural resource are driven to zero when the average return from the resource equals the marginal opportunity cost of exploiting an outside alternative. This is conceptually akin to, but not to be confused with, the term "rent seeking" which plays an important role in political economy and public choice. For the latter, see Tullock (1980).

⁷ This specification actually has a number of other possible interpretations. For instance, if one defines $F(\sum x_j)/\sum x_j = y$, 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).

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

$$-c + F'(\sum x_i) = 0. \quad (3)$$

According to (3), 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 (3) represents the maximal rent that can be extracted from the resource in a single period. This theory has largely been rejected by political economists and replaced by two noncooperative theories, each based on the assumption that no appropriator group will form and appropriators follow individually rational investment decisions.

One noncooperative theory (T2), due initially to Gordon (1954), asserts that the average group return will be driven down to the outside opportunity cost, as in (4):

$$c = F(\sum x_i) / \sum x_i. \quad (4)$$

According to this theory, all the potential rents from the CPR are fully dissipated, and each individual's return from the resource is driven down to cw , precisely what they would get if there were no resource. This theory is popularly known as the tragedy of the commons (Hardin, 1968) and is often associated with "open access" CPRs.

Another noncooperative theory (T3) takes the payoffs (1) to be the payoff function in a symmetric noncooperative game and then solves the game played by the appropriators

for a Nash equilibrium. Since our experimental design is symmetric, there exists a symmetric Nash equilibrium, where each player invests x_i^* in the CPR, satisfying (5):

$$-c + (1/n)F'(nx_i^*) + F(nx_i^*)((n-1)/nx_i^*) = 0. \quad (5)$$

At the symmetric Nash equilibrium, group investment in the CPR is greater than that under T1, but less than that under T2. Not all rents are dissipated. This is the noncooperative theoretical solution generally associated with a "limited access" CPR (see, for example, Clark, 1980; Comes and Sandier, 1986; and Negri, 1989).⁸

A final theory, based on the "Folk Theorem" for infinitely repeated games, assumes that individuals adopt trigger strategies. This theory predicts each individual plays according to the theory T1, unless a deviation occurs. If a deviation occurs, each individual plays a punishment strategy (invest according to T3) forever. Since cooperation supported by this trigger strategy is an equilibrium, there is no incentive to deviate. Indeed, this trigger strategy constitutes a subgame perfect equilibrium of the repeated game.

We have examined these four theories in the laboratory with eight appropriators ($n = 8$) and quadratic production functions $F(\sum x_i)$, where:

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

$$\text{with } F'(0) = a > c \text{ and } F'(nw) = a - 2bnw < 0.$$

For this quadratic specification, one has from (3) that the group optimal investment

⁸ 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, it is full dissipation (T2) which is predicted by non-zero consistent conjectures. See Mason, et al. (1988) for a discussion of consistent conjectures equilibria for the CPR experiment.

satisfies $\sum x_i = (a-c)/2b$. From (4), full rent dissipation occurs when investment in the CPR satisfies $\sum x_i = (a-c)/b$. Thus, rent is fully dissipated when appropriators invest twice as much in the CPR as is optimal. Finally, from (5), the symmetric Nash equilibrium group investment is given by:

$$\sum x_i = ((n-1)/n)(a-c)/b.$$

This level of investment is between maximal rent earnings and rent dissipation, and approaches the latter as n gets large. One additional constraint that arises in a laboratory setting is that the x_i be integer-valued. This is accomplished by choosing the parameters a , b , c , and w in such a way that the predictions associated with T1, T2, and T3 are integer valued.

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 PLATO (NOVANET) computer system at IU. The computer facilitates the accounting procedures involved in the experiment, enhances across experimental control, and allows for minimal experimenter interaction.

The Choice Environment

At the beginning of each experimental session, subjects were told that they would be making a series of investment decisions, that all individual investment decisions were anonymous to the group, and that at the end of the experiment they would be paid privately (in cash) their individual earnings. Subjects then proceeded to go through, at their own pace, a set of instructions that described the investment decisions.⁹

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 that 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 1 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 actual number of investment decision periods. All subjects were experienced, i.e. had participated in at least one

⁹ 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.

experiment using this form of decision environment.

Experimental Design

In the three "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 players' individual investment decisions. The parameters used in the baseline experiments are reported in Table I.

Given the parameters chosen for our baseline experiments, the solutions for T1, T2, and T3 lead to predictions displayed in Figure 2. A group investment of 36 tokens yields the Pareto optimal level of investment at which $MRP = MC$ and thus yields maximum rents (denoted T1). Conversely, a group investment of 72 tokens yields a level of investment at which $ARP = MC$ and thus zero rents from Market 2 (denoted T2). Finally, this symmetric game has a unique Nash equilibrium with each subject investing 8 tokens in Market 2 (denoted T3 in Figure 2). At the Nash equilibrium, subjects earn approximately 39.5 percent of maximum rents.¹⁰

Baseline Results

The baseline results are summarized in Table II and Figure 3. Aggregating across all experimental periods, the average level of rents accrued equalled -3.16 percent (28.6

¹⁰ See Walker, Gardner, and Ostrom (1990) for details of this derivation. Rents accrued as a percentage of maximum = (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.

percent over the last 10 periods). The percentage of optimal rents earned in each experiment (and the means) for each of the 20 decision rounds are presented in Figure 3.

Several characteristics of the individual experiments are important. In Figure 3, we see a pattern that (to varying degrees) we found across all three experiments. Investments in Market 2 show a "pulsing" pattern in which rent is reduced, at which time investors tend to reduce their investments in Market 2 and rents 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 "rent peaks."¹¹ Note, for these experiments, the low points in the pulsing pattern were at rent levels far below zero. Over the course of the experiments, there was a tendency for the variance in rents to decrease and average rents to increase. We did not see, however, any clear signs that the experiments were stabilizing or approaching full efficiency. Further, we observed no experiments in which the pattern of individual investments in Market 2 stabilized at the Nash equilibrium. This is a behavioral result we have found consistently across over 50 experimental replications with alternative parameterizations.¹² At the aggregate level, the Nash prediction (T3) 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, individuals acting independently in a CPR situation do not come close to solving the collective action problem they confront. We now turn to the sanctioning problem.

¹¹ This pulsing pattern is similar to the patterns of investment observed in numerous other experiments we report in Gardner, et al. (1990) and Walker, et al. (1990).

¹² Several of our parameterizations explicitly examine whether the magnitude of the separation of payoffs at the optimum and at the Nash equilibrium affects behavior. We have examined "high" pay and low "pay" designs and one design in which the MC (return from Market 1) was zero. These changes in parameterizations do not have a significant impact on behavior.

IV. SANCTIONING EXPERIMENTS: Experimental Design, Theory, and Results

Experimental Design I

In field settings, appropriators must invest resources in monitoring each other's activities as well as in taking costly actions to sanction one another. In our initial experimental settings, we wished to focus entirely on sanctioning without the additional complication of the cost of monitoring. Consequently, we provided free, but anonymous, information about the activities of all participants. Thus, all Design I sanctioning experiments began just like baseline experiments with the exception that after each decision round, subjects also received information on the individual decisions of all players. The individual information was given by subject number, with each subject's identity thereby masked.

Our sanctioning mechanism requires that each individual incurs a cost (a fee) in order to sanction another individual. 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 the Design I experiments, after 10 decision rounds, the subjects were asked to read an announcement which is summarized below (see APPENDIX A for the actual announcement).

The subjects were informed that in all remaining periods they would be given the opportunity to place a "fine" on another player in the experiment and/or have other players place a "fine" on them. Each player could levy one fine at a specified fee. The player fined would pay a fine which was double the fee charged to fine another player. Thus, players could place only one fine, but it was possible for a single player to be charged with multiple fines. After each round, each player was asked

to fill out a "FINE SHEET." The fine sheets were collected privately by the experimenter. The experimenters tallied all fines and reported the results back to each player. Note that any player who was fined did not know the identity of the other player or players who imposed the fine. At no time during or after the experiment did players know the actual identity of other players, nor did they know which players chose to impose fines or which players were fined by other players. 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 which were used are reported in Table III. 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". The sanctioning mechanism was in place for up to 15 periods. Subjects did not know in advance the end period of the experiment.

Theory: Perfect and Imperfect Equilibrium with Sanctioning

First, we show how the possibility of sanctioning extends the strategy space of the CPR game, and how repetition of the one-shot game with sanctioning leads to a unique subgame perfect equilibrium. The investment levels, according to this subgame perfect equilibrium, are the same as those of T3 (the Nash equilibrium without sanctioning) for the one-shot game. Second, we show the existence of imperfect equilibria when players take advantage of richer threat possibilities than those afforded by subgame perfection. Indeed, these imperfect equilibria exhibit significantly more sanctioning activity than predicted by trembles and considerably less rent dissipation than predicted by subgame perfection.

The sanctioning institution is represented formally using the following construction. Let Y be a matrix of 0's and 1's, where $y_{ij} = 1$ means that player i has sanctioned player j , and $y_{ij} = 0$ means that i has not sanctioned j . Row i of the matrix Y 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, $u_i(x, Y)$, is given by:

$$u_i(x, Y) = u_i(x) - f_1 \sum_j y_{ij} - f_2 \sum_j y_{ji}. \quad (6)$$

The parameters f_1 and f_2 represent the cost of fining and the cost of being fined, respectively. The sum $\sum_j y_{ij}$ is the total number of fines j levied by player 1, costing him f_1 each; the sum $\sum_j y_{ji}$ is the total number of times player i is fined, costing him f_2 each.

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 $(x^*, Y^*) = (x^*, 0)$. That is, 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 (7):

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

In the event of any deviation from prescribed play,
resume playing $(x^*, 0)$ after the deviation. (7)

This equilibrium follows from backward induction. At the last period T , no deviation is profitable. At the next to last period $T-1$, 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 which predicts subgame perfect equilibrium play, no sanctioning will take

place (except perhaps by mistake) and rent dissipation will be the same as in the one-shot Nash equilibrium case. This conclusion continues to hold even if players do not know exactly when the game will end (T is a random variable), but they know that the game will end in a finite amount of time. Suppose the most that T could be is 30 periods. Then if period 30 is reached, they would play the unique equilibrium at time $T=30$. 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 (8):

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

In the event of any deviation, play $(x_i = 25, Y = 1)$ for one period, then resume playing $(z,0)$. (8)

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 (8). We claim that (8) represents an imperfect equilibrium. To show this, it suffices to show that no deviation from prescribed play pays. For our production function, $F(200)$ is a very large negative number. A player who deviated optimally for one period would gain at most \$.50, depending on the level of z_i , but in the next period would lose \$4.20 due to punishment from overinvestment alone, plus $f_1 + f_2$ due to all players sanctioning and being sanctioned, as in (8). 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 (8) is an equilibrium. The imperfection of (8) lies in the fact that the trigger punishment -- dumping all tokens into the resource, everybody fining himself -- 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 $Y = I$, the identity matrix, so that everyone sanctions himself 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 identify 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 (8) will generate n fines every time a mistake takes place, which is considerably more than the 0 fines generated by the subgame perfect equilibrium (7).

There is a large set of equilibria along the lines of (8), 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 f_1 and f_2 , we had hoped to allow the subjects to find equilibria of the family (8) 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.

Design I Results:

The principle results from our Design I experiments can be summarized as follows:

- (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 Market 2 investors;
- (3) there is a nontrivial amount of sanctioning which can be classified as error, lagged punishment, or "blind" revenge;
- (4) independent sanctioning as employed in this experimental context has a modest impact on rent dissipation; and
- (5) when fining fees and sanctioning costs are included in measurements of efficiency, the gains in efficiency due to lower rent dissipation are wiped out.

Our first principle finding is that subjects actually sanctioned each other at a much higher rate than the zero rate predicted by subgame perfection (see Table IV). We observed 176 instances of sanctioning across the eight experiments. In no experiment did we observe fewer than 10 instances of sanctioning. The frequency of sanctions is inversely related to the cost of imposing the fine and dramatically increases with the stiffness of the fine. Further, our results, although reminiscent of equation (8), do not strictly support the conclusion that players were playing an equilibrium of this form. Except for experiment 79, where the degree of rent dissipation was less than 5%, our experiments reveal patterns of rent dissipation and levels of sanctioning which are too inefficient to be 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 Market 2 in either the no sanctioning periods or in earlier rounds of the sanctioning periods;
- (c) Blind revenge -- the person fined was a low Market 2 investor and was fined by a person fined in a previous period; and
- (d) Error — no obvious explanation can be given for the action (trembling hand).

In Table IV, we summarize the frequency of fines falling in each category [we have combined (3) and (4) due to low frequency and difficulty in distinguishing between the two]. Several conclusions can be drawn from this very preliminary analysis. Seventy-seven percent of all sanctioning is aimed at investors who in the previous period were above-average investors in Market 2. An additional 7 percent were aimed at players who had been heavy investors in Market 2 in earlier (but not the most recent) periods. We would classify an additional 5 percent in the blind revenge category and the remaining 11 percent as errors.

The fourth result focuses on the level of rent dissipation when sanctioning is imposed. The results from the eight experiments in which the sanctioning mechanism was used are reported in Figure 4. This figure contrasts the mean level of rents accrued as a percentage of optimum in the baseline experiments with the results from the experiments with sanctioning. Clearly, the first 10 periods (when no sanctioning was in place) does not significantly differ between the two sets of experiments. This result suggests that the addition of anonymous information on individual decisions had no observable impact on investments. In periods 11-20, one might argue that sanctioning had some effect in

increasing the level of rents earned and thus increasing efficiency. However, the effect on the level of rent accrual is modest: average rents over periods 11-20 were 28.6% in the baseline and 39.4% in experiments with sanctioning.

In Figures 5a-5h, we report across period behavior within each of the sanctioning experiments. In the top panel of each figure the level of rents accrued as a percentage of optimum is displayed. The lower panel summarizes period by period information on the number of fines which were placed each period and (above each bar) the number of individuals that were fined. Several features of these experiments stand out. In only one of the eight experiments do we see the sanctioning mechanism having a major impact on the level of rents (Experiment E79 - 20/80). In this experiment, where the fee to fine ratio was 20/80, the level of rents increased from an average of -60% in the first 10 periods to 95.2% in the final 15 periods. Fines were placed frequently (an average of 2.6 per period) and primarily on two subjects (28 of 37 fines) who tended to be relatively high Market 2 investors.

Our final result incorporates gains in efficiency with costs incurred with sanctioning. In summary, fining fees and sanctioning costs offset gains in efficiency arising from lower rent dissipation (decreased investment in Market 2). In Table V, we include fining fees and sanctioning costs into calculations of collective benefits from participating in the experiment. Specifically, we define a measurement of efficiency where actual experimental payoffs are measured as a percentage of maximum obtainable payoffs. When one takes into account sanctioning costs and fining fees, one observes lower levels of efficiency in independent sanctioning experiments than in the baseline. Comparing the sanctioning

periods (11-20+) in our independent sanctioning experiments to periods (11-20) from the baseline experiments we find a 3 percent drop in average efficiency. Experiment 79 is an interesting case. During the sanctioning phase, rents averaged almost 100 percent of optimum. Taking into account fees and fines, efficiency was reduced to 80 percent of optimum.

Experimental Design II

The Design II experiments were conducted as a check on the robustness of our original design. Readers of an earlier version of this manuscript conjectured that the lack of a significant improvement in rent accrual with the introduction of a sanctioning mechanism in Design I could be due to a hysteresis effect tied to the decisions in the first 10 periods, periods in which there was no sanctioning mechanism. With Design II, 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. The results from these three new experiments are summarized in Figure 6. Plotted is the mean level of rent accrual for each decision period. These measures are contrasted with the means from our baseline experiments in which no sanctioning mechanism was available to subjects. As one can see, the results are consistent with those for our Design I experiments. There is 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.

VI. CONCLUSIONS

In an experimental CPR setting, we find that individuals independently impose costly sanctions when subgame perfect equilibrium theory predicts they will not. Sanctioning is primarily directed toward individuals who heavily exploit the CPR. No evidence of trigger strategies is found. Even though sanctioning is at a substantial level, rent dissipation continues to be a serious problem.. Further, when fining fees and sanctioning costs are accounted for, sanctioning does not lead to an improvement in efficiency.

In some field settings, appropriators sanction to enforce self-negotiated rules regarding CPR access. There is considerable evidence that such sanctioning does enhance efficiency. In previous experimental research, we have found that allowing appropriators to communicate on a regular basis, even when investment decisions are made anonymously, enables them to achieve near optimal rents (Ostrom and Walker, 1991). In ongoing research, we examine the joint effect of communication and endogenous choice of a sanctioning institution on efficiency. Communication gives players the opportunity to negotiate contractual arrangements via coordinated strategies. Preliminary results suggest that this treatment substantially enhances overall efficiency. These results leave unanswered the question of why communication in conjunction with endogenous choice of a sanctioning mechanism works so well in without external enforcement. This is one of the primary issues to be addressed in our future field and experimental research.

APPENDIX A

SPECIAL 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 times \$.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 a 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 = $(1/2) \times \$15.50 =$	\$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 WILL OCCUR AFTER THE INVESTMENT DECISIONS ARE REPORTED FOR THE NEXT PERIOD (PERIOD 11).

Bibliography

- Acheson, James M. (1989) "Where Have All the Exploiters Gone? Co-Management of the Maine Lobster Industry." In Fikret Berkes, ed. Common Property Resources. London: Belhaven Press, 199-217.
- Alexander, P. (1982) Sri Lankan Fisherman: Rural Capitalism and Peasant Society. Australian National University Monographs on South Asia No. 7. Canberra: Australian National University.
- Arnold, J.E.M. and J. G. Campbell. (1986) Collective Management of Hill Forests in Nepal: The Community Forestry Development Project. In National Research Council, Proceedings of the Conference on Common Property Resource Management. Washington, D.C.: National Academy Press, 425-454.
- Baines, Graham B. K. (1989) "Traditional Resource Management in the Melanesian South Pacific: A Development Dilemma." in Fikret Berkes (ed.) Common Property Resources. London: Belhaven Press, 273-295.
- Bell, Frederick W. (1972) "Technological Externalities and Common Property Resources: An Empirical Study of the U.S. Lobster Industry." Journal of Political Economy. Vol. 80: 148-158.
- 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.
- 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 P. T. Liu, ed. Dynamic Optimization and Mathematical Economics. New York: Plenum Press, 117-132.
- Cordell, John C. and Margaret A. McKean. (1986) "Sea Tenure in Bahia, Brazil." In National Research Council, Proceedings of the Conference on Common Property Resource Management. Washington, D.C.: National Academy Press, 85-113.
- Comes, 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 National Research Council, Proceedings of the Conference on Common Property Resource Management, Washington, D.C.: National Academy Press, 115-135.
- 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 C. Lamson and A. J. Hanson, eds. Atlantic Fisheries and Coast Communities: Fisheries Decision-Making Case Studies. Halifax, Canada: Dalhousie Ocean Studies Programme.
- Feeny, D. H. (1988) "Agricultural Expansion and Forest Depletion in Thailand, 1900-1975." In John F. Richards and Richard P. Tucker, eds. World Deforestation in the Twentieth Century. Durham, North Carolina: Duke University Press, 112-143.
- Frohlich, Norman and Joe Oppenheimer. (1970) "I Get By With a Little Help From My Fields." World Politics, Vol. 23: 104-120.
- Gadgil, Madhav and Prema Iyer. (1989) "On the Diversification of Common-Property Resource Use by Indiana Society." In F. Berkes, ed. Common Property Resources. London: Belhaven Press, 240-272.
- Gardner, Roy, Elinor Ostrom, and James Walker (1990) "The Nature of Common-Pool Resource Problems." Rationality and Society, Vol. 2: 335-358.
- Gordon, Scott (1954) "The Economic Theory of a Common Property Resource: The Fishery," J. Pol. Economy, 124-142.
- Hardin, Garrett. (1968) "The Tragedy of the Commons." Science, Vol. 162: 1243-1248.
- Isaac, Mark and James Walker (1988) "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism," Quarterly J. Economics, February: 179-200.
- Ledyard, John O. (1990) "Is there a Problem with Public Good Provision." Pasadena: California Institute of Technology. Mimeo.
- Libecap, Gary D. (1990) Contracting for Property Rights. Cambridge: Cambridge University Press.
- Mason, C, Todd Sandier, and Richard Comes (1988) "Expectations, The Commons, and Optimal Group Size," J. of Environmental Economics and Management, 15, 99-110.

- McHugh, J. L. (1972) "Jeffersonian Democracy and the Fisheries." In B. J. Rothschild, ed. World Fisheries Policy: Multidisciplinary Views. Seattle: University of Washington Press.
- McKean, Margaret (1986) "Management of Traditional Common Lands (Iraichi) in Japan." In Proceedings of the Conference on Common Property Resource Management, National Research Council. Washington, D.C.: National Academy Press, 533-89.
- Messerschmidt, Donald A. (1986) "Collective Management of Hill Forests in Nepal: The Community Forestry Development Project." In National Research Council, Proceedings of the Conference on Common Property Resource Management. Washington, D.C.: National Academy Press, 455-480.
- 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, Vol. 25: 9-15.
- Netting, Robert McC. (1981) Balancing on an Alp: Ecological Change and Continuity in a Swiss Mountain Community. Cambridge: Cambridge University Press.
- Ostrom, Elinor (1990) Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press.
- Ostrom, Elinor and James Walker. (1991) "Communication in a Commons: Cooperation without External Enforcement." Forthcoming Contemporary Laboratory Research in Political Economy, ed. Thomas Palfrey, University of Michigan Press.
- 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, Vol. 93: 106-127.
- Plott, Charles R. and Robert A. Meyer (1975) "The Technology of Public Goods, Externalities, and the Exclusion Principle." in Edwin S. Mills, (ed.) Economic Analysis of Environmental Problems. New York: National Bureau of Economic Research.

- 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.
- Selten, Reinhard (1971) "A Simple Model of Imperfect Competition Where 4 Are Few and 6 Are Many." International Journal of Game Theory. Vol. 2: 141-201.
- Siy, Robert Y. (1982) Community Resource Management: Lessons from the Zanjera. Quezon City, Philippines: University of the Philippines Press.
- Taylor, Michael. (1987) The Possibility of Cooperation. Cambridge: Cambridge University Press.
- Thomson, Jamie T. (1977) "Ecological Deterioration: Local-Level Rule Making and Enforcement Problems in Niger." In M. H. Glantz, ed. Decertification: Environmental Degradation in and Around Arid Lands. Boulder: Westview Press, 57-79.
- Thomson, Jamie T., David Feeny, and Ronald J. Oakerson. (1986) "Institutional Dynamics: The Evolution and Dissolution of Common Property Resource Management." In National Research Council, Proceedings of the Conference on Common Property Resource Management. Washington, D.C.: National Academy Press, 391-424.
- Tullock, Gordon (1980) "Efficient Rent-Seeking." In J. M. Buchanan, R. Tollison, and G. Tullock, eds. Towards a Theory of the Rent-Seeking Society. College Station: Texas A&M 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. Vol. 19: 203-211.

TABLE I

EXPERIMENTAL DESIGN BASELINE
Parameters for a Given Decision Period

Experiment Type:	25 Tokens
Number of Subjects	8
Individual Token Endowment	25
Production Function: Mkt.2*	$23(\sum x_i) - .25(\sum x_i)^2$
Market 2 Return/unit of output	\$.01
Market 1 Return/unit of output	\$.05
Earnings/Subject at Group Max.**	\$.83
Earnings/Subject at Nash Equil.	\$.70
Earnings/Subject at Zero Rent	\$.63

* $\sum x_i$ = 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.

** Subjects were paid in cash one-half of their PLATO earnings.
Amounts shown are potential cash payoffs.

TABLE II

DESCRIPTIVE STATISTICS: BASELINE
 PERCENTAGE OF MAXIMUM RENTS ACCRUED
 (mean - standard deviation - range)

1X25:	5.32	65.28	(-160 to 75)
2X25:	-28.22	106.40	(-382 to 97)
3X25:	13.41	50.97	(-109 to 63)
Pooled:	-3.16	78.64	(-382 to 97)

TABLE III

SANCTIONING FEE AND FINE PARAMETERS

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

TABLE IV

TABULATIONS BY REASON FOR SANCTIONING

EXPERIMENT # FEE/FINE	EXPS.52;53 5/10	EXP.56 5/20	EXPS.77;88 40/80	EXPS.79;83;84 20/80	EXP.AVE.
REASON:					
(1)	50	33	23	29	16.9
(2)	1	1	1	10	1.6
(3&4)	3	8	1	2	1.8
EXP. AVE.	27	42	13.5	17.7	

TABLE V
EFFICIENCY MEASUREMENTS
(ACTUAL EARNINGS/MAXIMUM POTENTIAL EARNINGS)*

		MEAN EFFICIENCY	
Experiment		First 10 Decision Periods	Decision Periods 11-20+
Baseline	E35	.735	.801
	E39	.613	.760
	E40	<u>.716</u>	<u>.860</u>
		Ave. = .688	Ave. = .807
Sanctioning	E52 (5/10)	.653	.747
	E53 (5/10)	.655	.798
	E56 (5/20)	.718	.767
	E77 (40/80)	.640	.753
	E78 (40/80)	.520	.785
	E79 (20/80)	.609	.797
	E83 (20/80)	.697	.761
	E84 (20/80)	<u>.813</u>	<u>.808</u>
	Ave. = .661	Ave = .777	

* Includes fines and fees paid.

Sanctioning experiments had no sanctioning option in periods 1-10.
Sanctioning followed each period after period 11 - experiments were conducted from 21 to 25 decision periods.

FIGURE 1

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

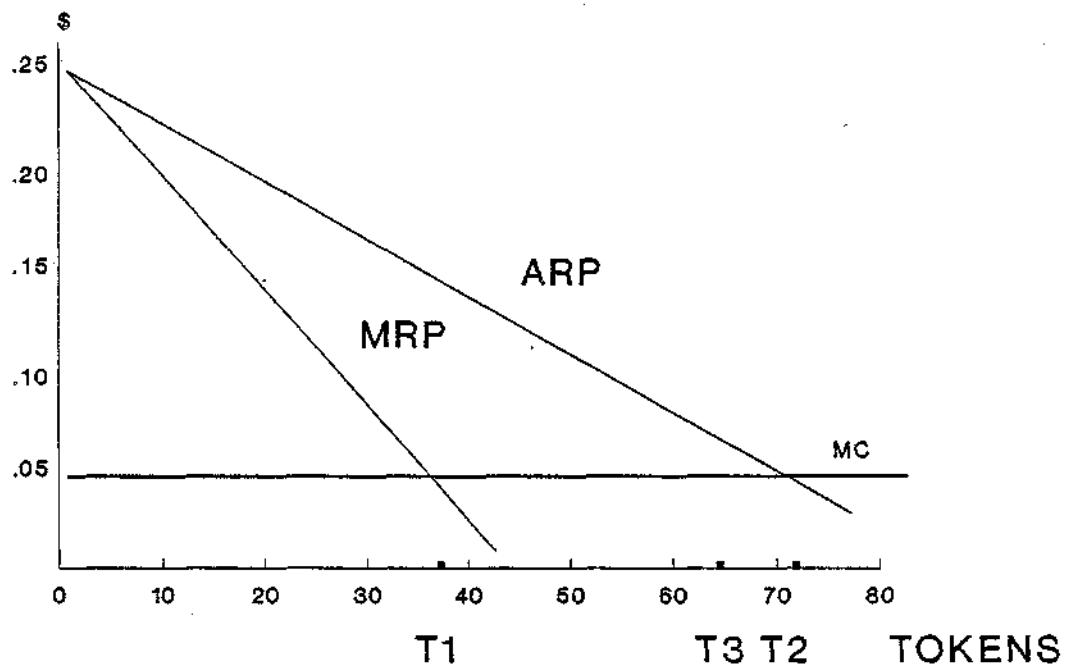
Tokens Invested by Group	Units of Commodity 2 Produced	Total Group Return	Average Return per Token	Additional Return per Token
20	360	\$ 3.60	\$ 0.18	\$ 0.18
40	520	\$ 5.20	\$ 0.13	\$ 0.08
60	480	\$ 4.80	\$ 0.08	\$-0.02
80	240	\$ 2.40	\$ 0.03	\$-0.12
100	-200	\$ -2.00	\$-0.02	\$-0.22
120	-840	\$ -8.40	\$-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

FIGURE 2

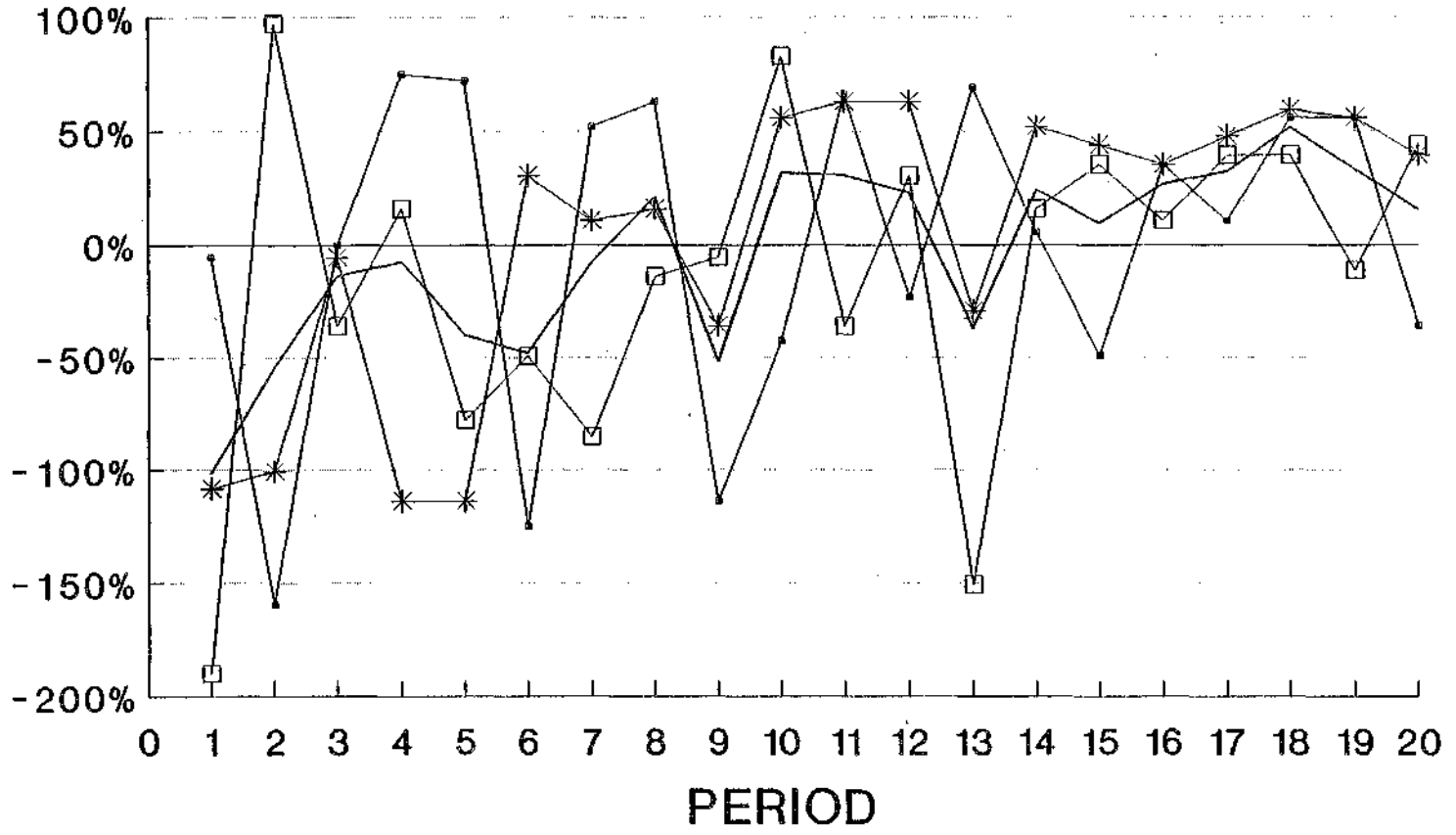
THEORETICAL PREDICTIONS



PREDICTIONS: T1=36, T2=72, T3=64

BASELINE EXPERIMENTS

RENTS AS A PERCENTAGE OF OPTIMUM



—•— SERIES 1 -□- SERIES 2 *— SERIES 3 — MEAN SERIES

* Exp.2 (period 1) = -382%

FIGURE 3

SUMMARY RESULTS

BASELINE VERSUS SANCTIONING: DESIGN I

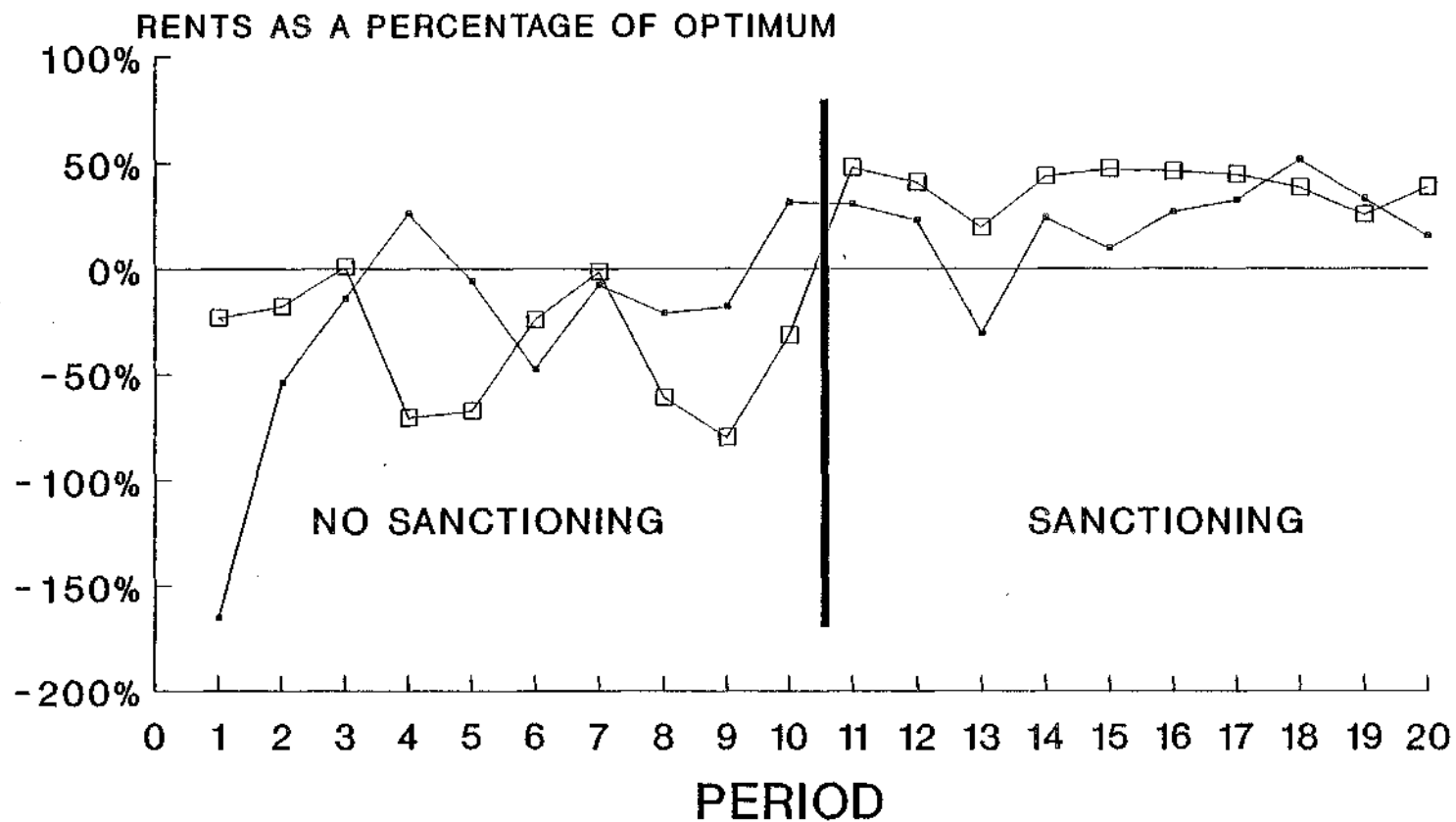


FIGURE 4

FIGURE 5a

EXPERIMENT E52 - 5/10

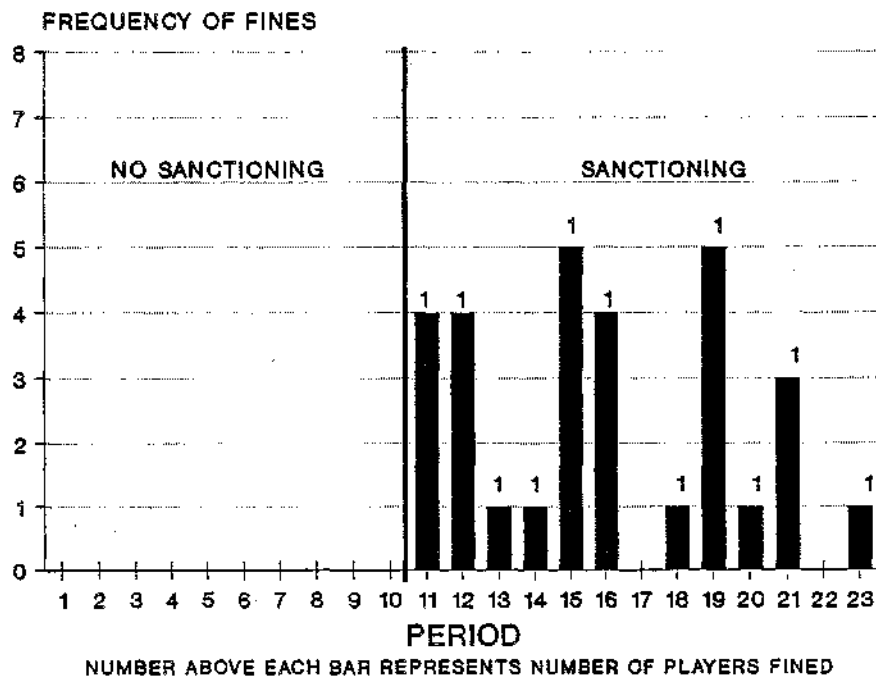
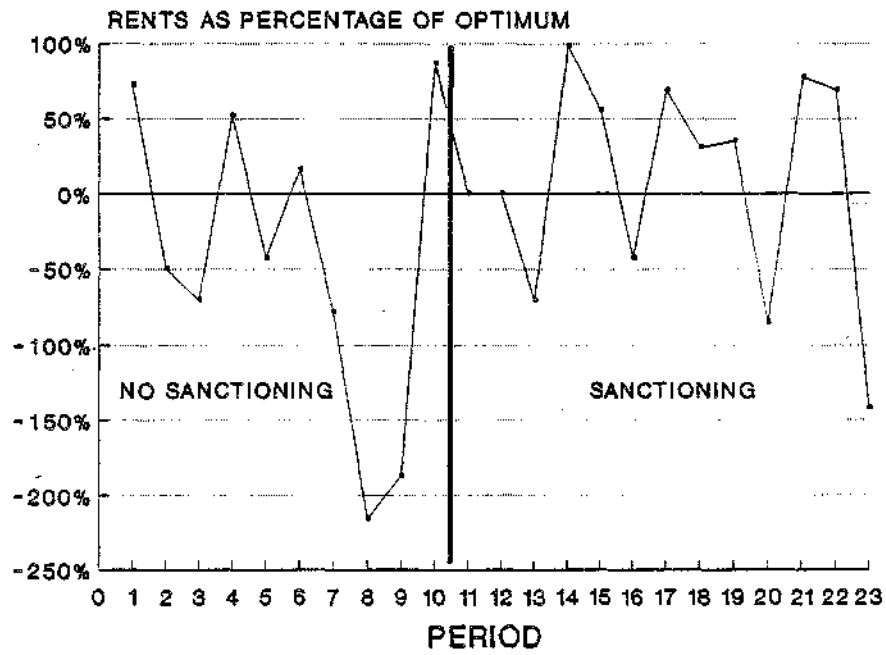


FIGURE 5b

EXPERIMENT E53 - 5/10

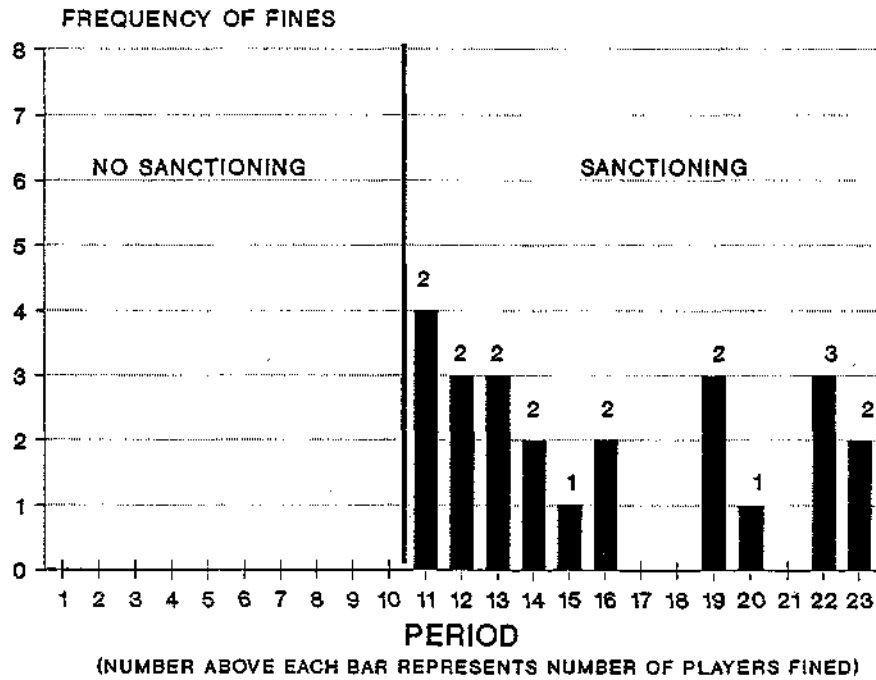
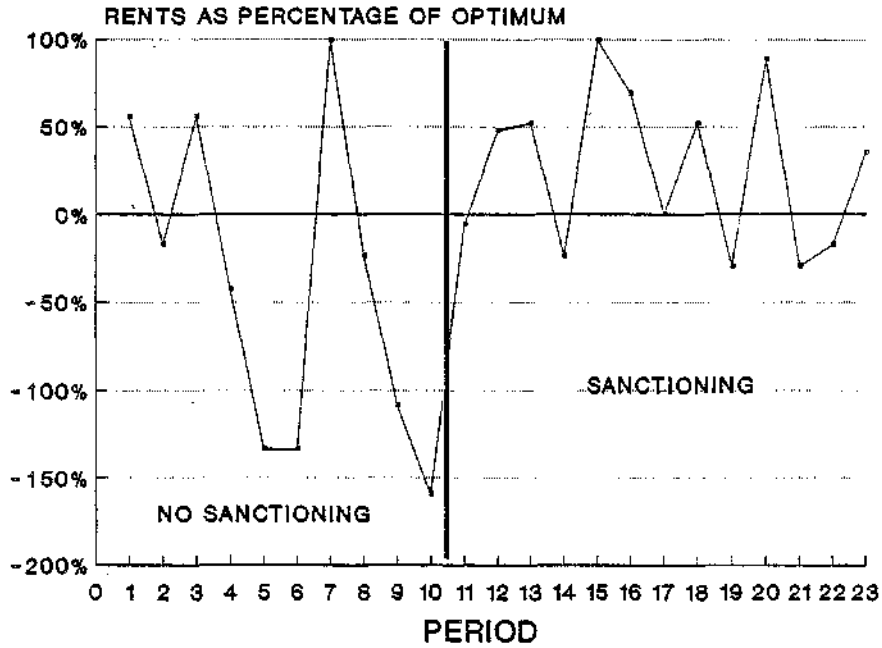


FIGURE 5c

EXPERIMENT E56 - 5/20

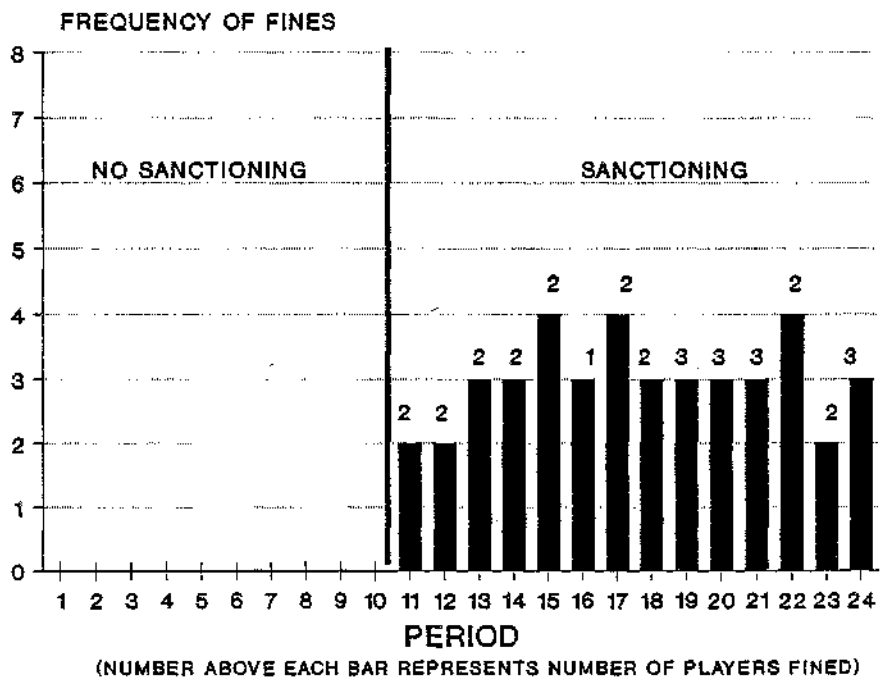
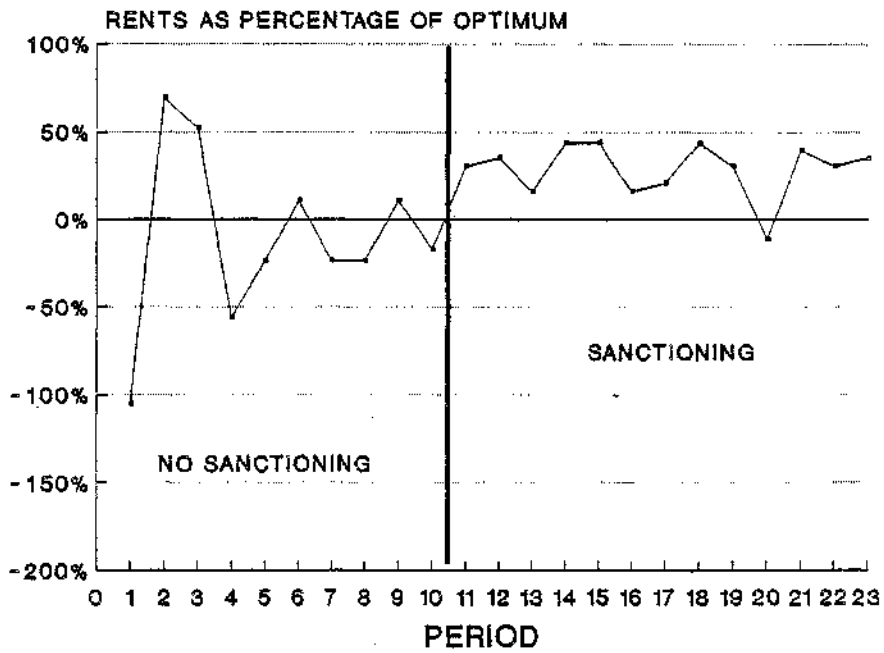


FIGURE 5d

EXPERIMENT E77 - 40/80

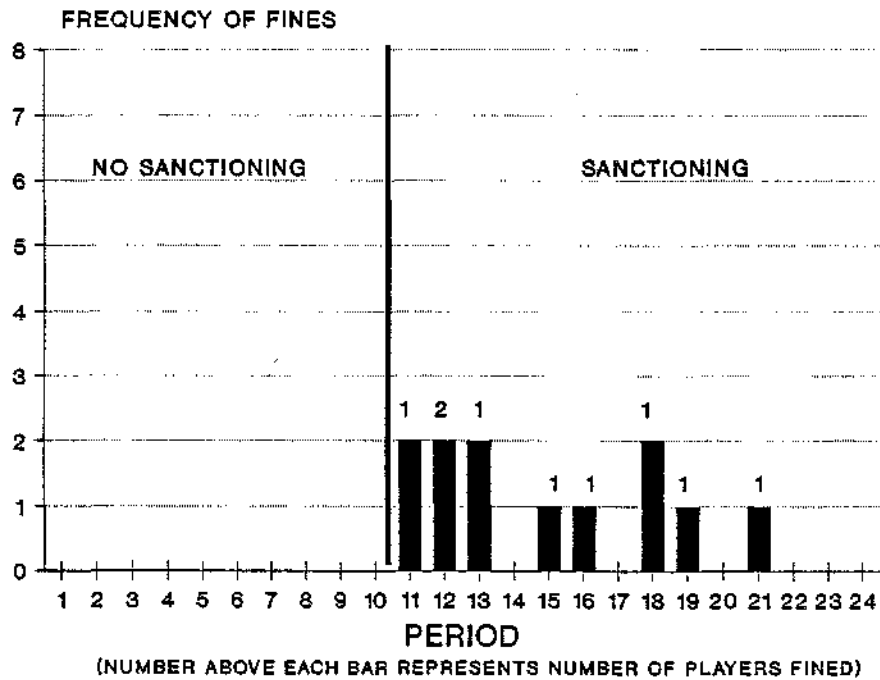
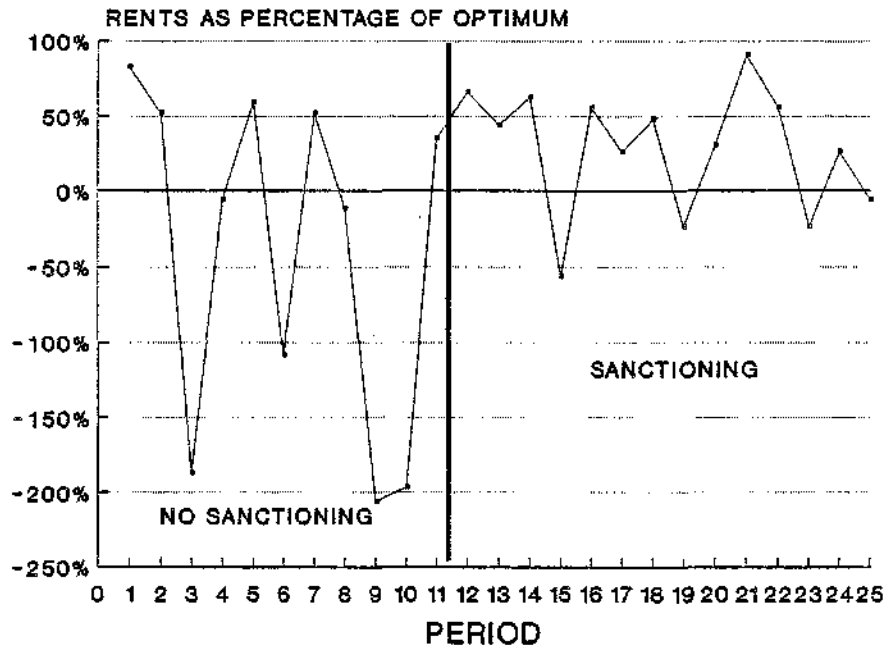


FIGURE 5e

EXPERIMENT E78 - 40/80

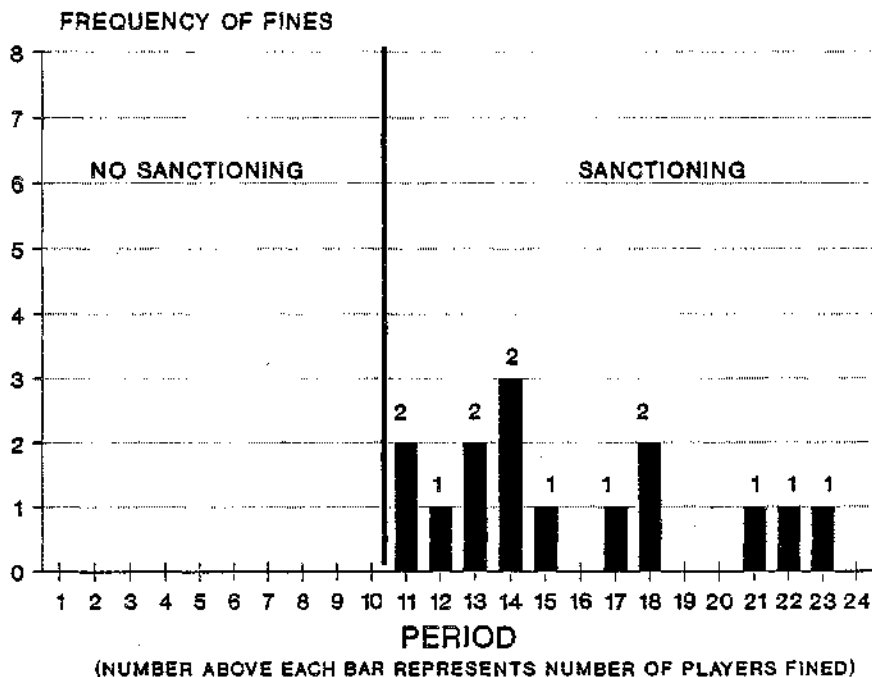
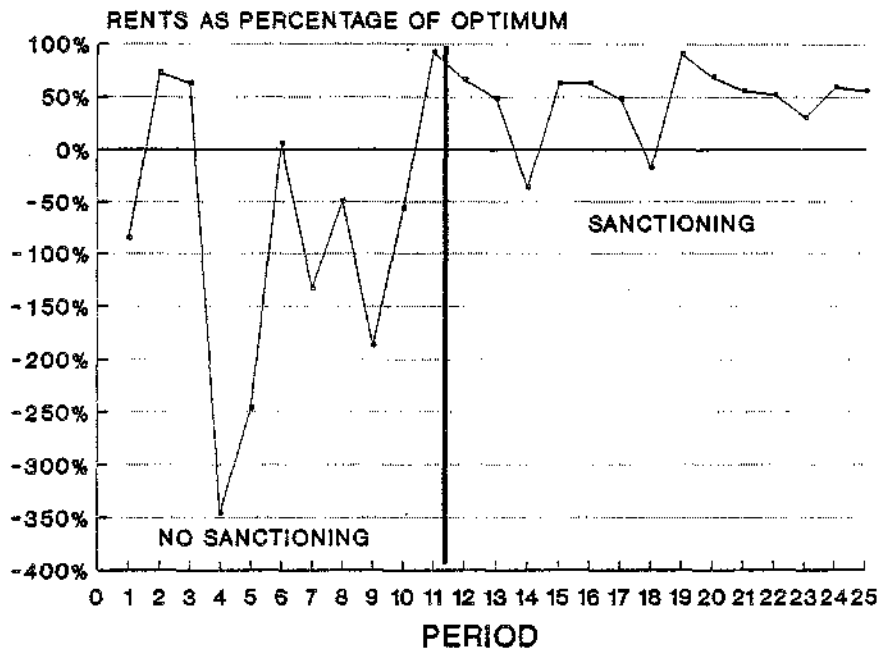


FIGURE 5f

EXPERIMENT E79 - 20/80

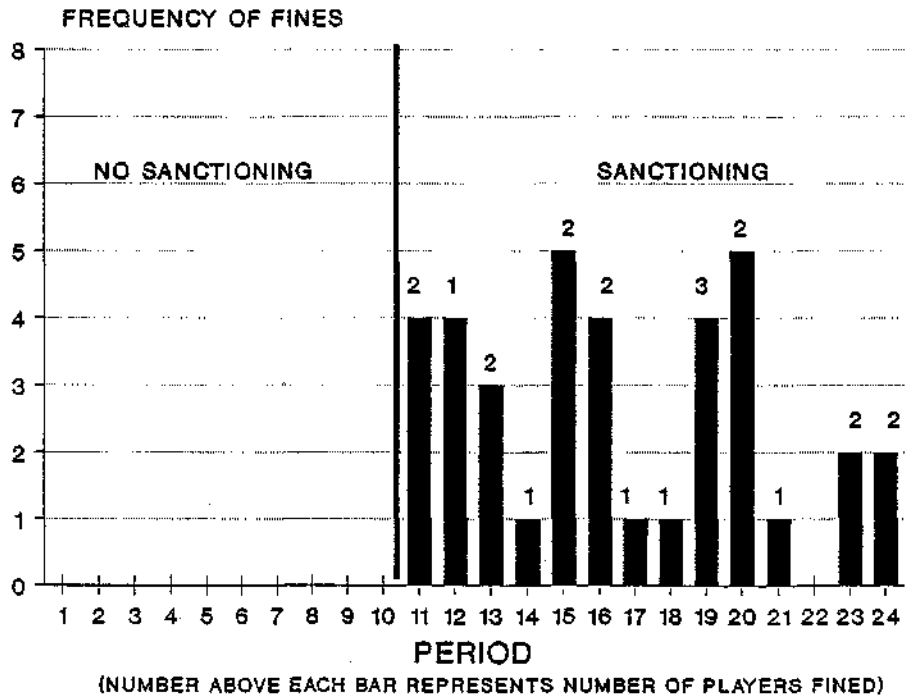
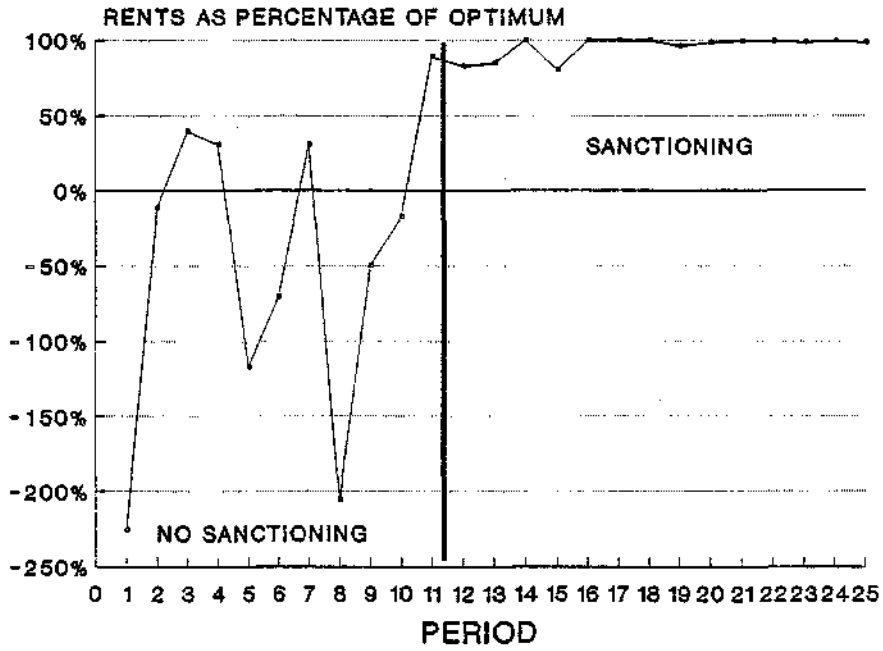


FIGURE 5g

EXPERIMENT E83 - 20/80

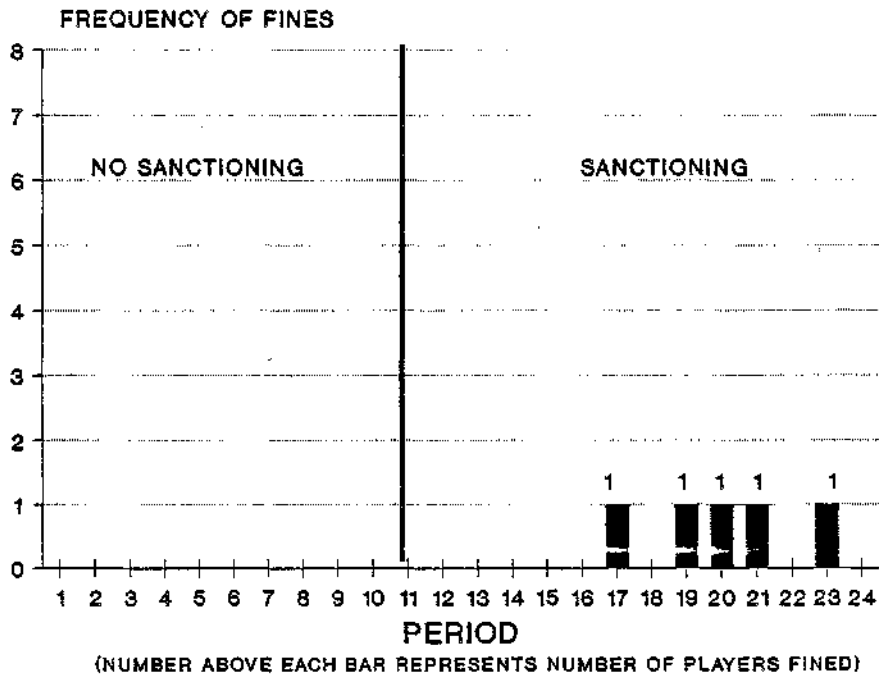
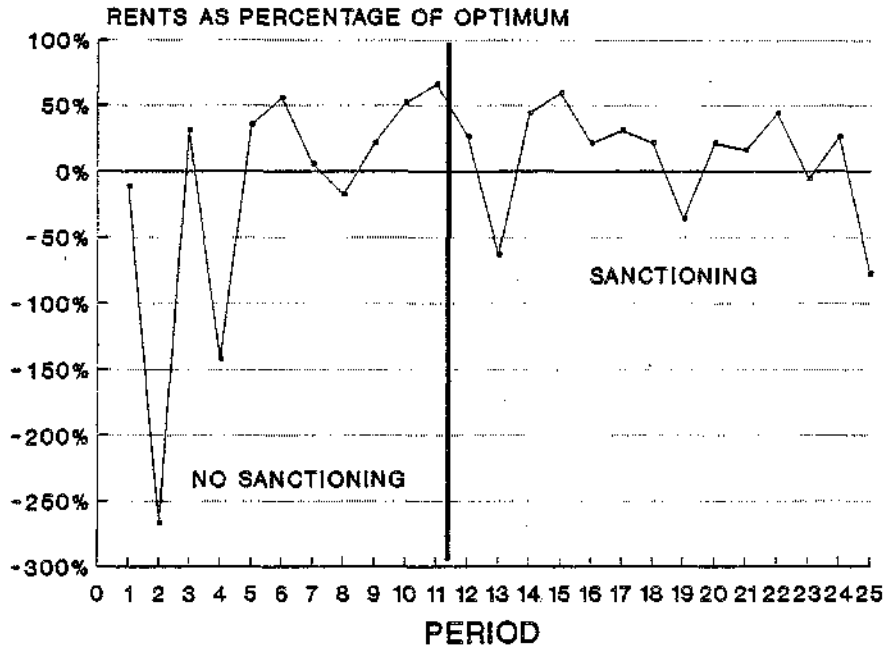
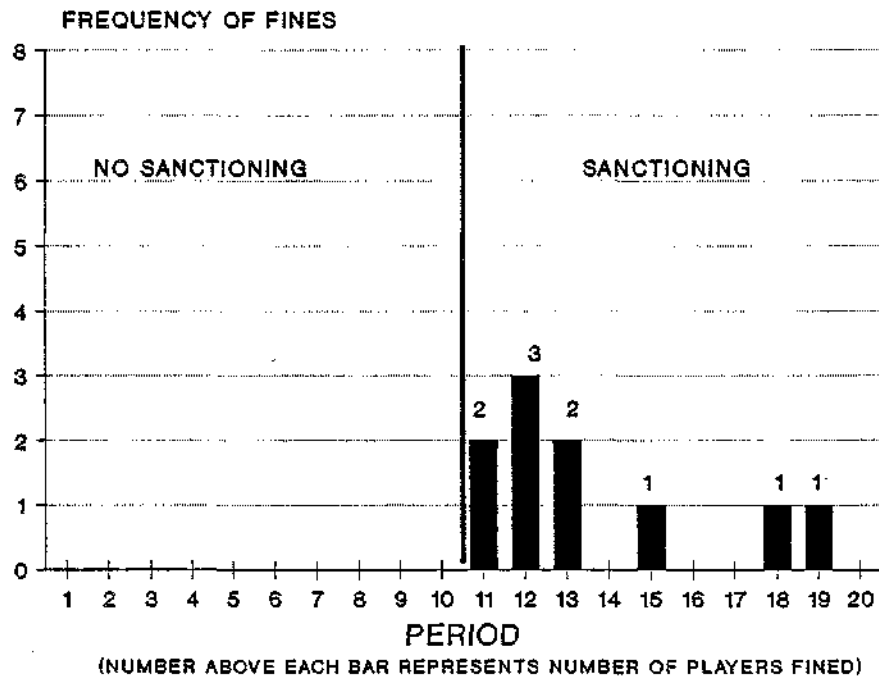
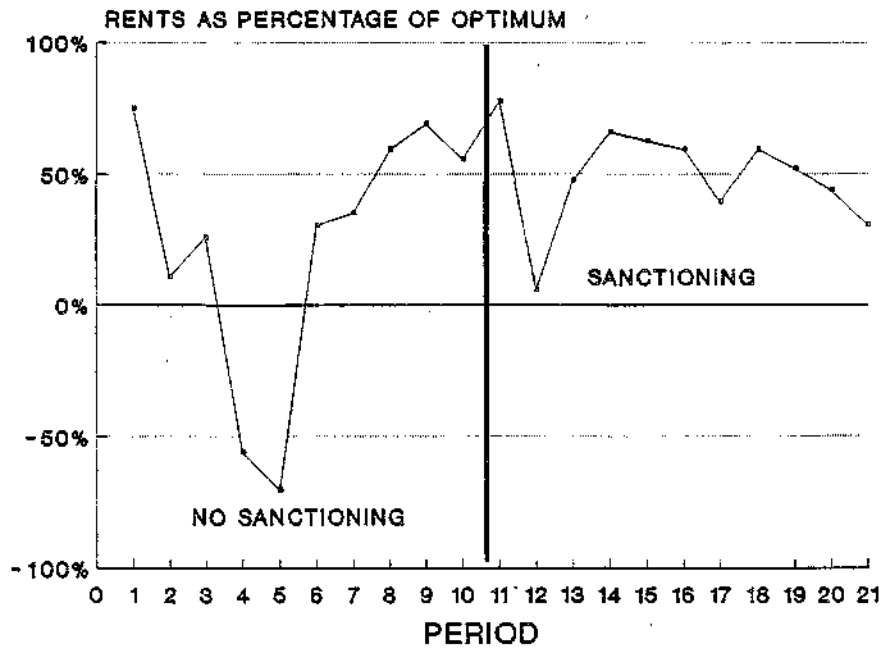


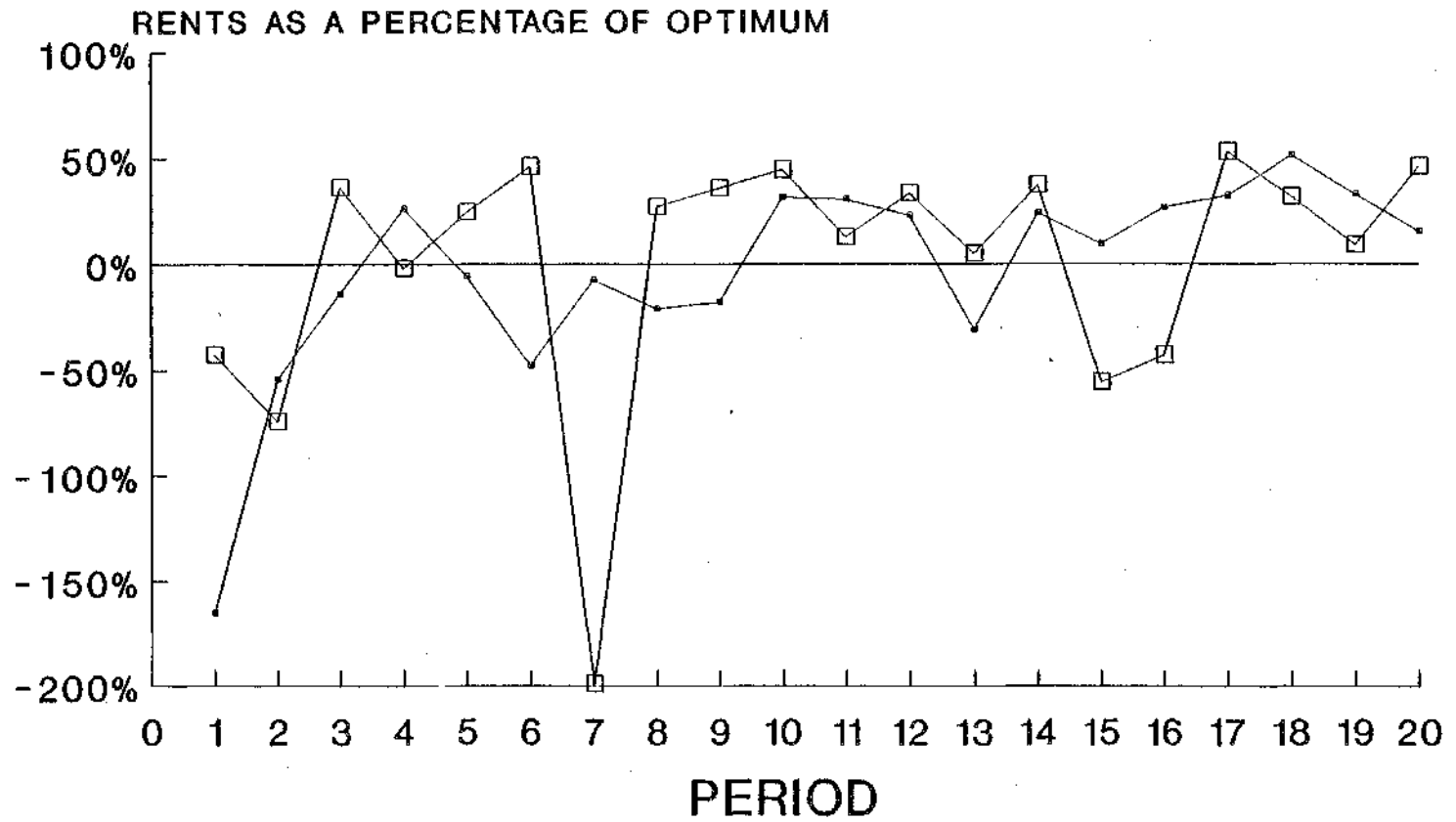
FIGURE 5h

EXPERIMENT E84 - 20/80



SUMMARY RESULTS

BASELINE VERSUS SANCTIONING: DESIGN II



—•— BASELINE —□— SANCTIONING

FIGURE 6