KEEPING AN EYE ON YOUR NEIGHBORS: AGENTS MONITORING AND SANCTIONING ONE ANOTHER IN A COMMON-POOL RESOURCE ENVIRONMENT

Marco Casari and Charles R. Plott[1]

Division of the Humanities and Social Sciences, 228 -77
California Institute of Technology
Pasadena, CA 91125
marco @ hss.caltech.edu   cplott @ hss.caltech.edu

March 2000

Abstract

The role of a specific institution in avoiding a "tragedy of the commons" situation in a common pool-resource environment is studied experimentally. The resource users privately decide their own exploitation level and then, once the group outcome is revealed, can choose to select other individuals for inspection. At a cost the inspector can view the decision of any individual. If the inspected individual has exploited the resource excessively, relative to a publicly known amount, a fine is imposed and paid to the inspector. The rules were modeled after an historical case of self-governed rural communities.

The introduction of the sanctioning institution greatly improves the efficiency of the group outcome from the initial level of severe "tragedy". The classical model with homogeneous, self-interested agents cannot explain these results. We present a model with heterogeneous, other-regarding agents that is compatible with both the resource use and the inspection decision patterns. In particular, differences in altruism/spite can explain the wide diversity of individual behavior and the willingness of spiteful agents to request unprofitable inspections help explaining the high inspection rate.

---

[1] The financial support of the National Science Foundation and the Laboratory of Experimental Economics and Political Science is gratefully acknowledged. We are also deeply appreciative of the collaboration of Stephen van Hooser who developed the software program and helped extensively with the experiments. Many thanks to Anthony Kwasnica, Alvaro Gonzalez Staffa, Leslie Title, Roberto Weber, Angela Hung, and Peter Coughlan for their suggestions and help in running the experiments. Thanks also to William Morrison for the comments, James Walker for providing a copy of the instructions of their experiment, and to Robert Moir for reviewing Table 1. An earlier version of this paper has been presented at the ESA annual meeting on May 28-30, 1999 and at a seminar at the University of Trento, Italy.

1. INTRODUCTION

Can we stimulate cooperative behavior in a group by allowing agents to monitor one another's actions and to sanction any free rider that is discovered? This paper explores such ideas in a series of laboratory experiments in a context where cooperation has proven to be especially challenging to reach, namely, in a common-pool resource environment.

The possibility of monitoring individual actions at a cost and of inflicting a punishment whenever a violation of an agreed upon rule occurs is frequently observed in reality as a way to enforce cooperation. We focus our attention on decentralized mechanisms where inspecting decisions are taken by the same agents who have the choice to cooperate or free ride (self-administered institution). An historical example of forest and pasture management in the Italian Alps (Casari, 1997) provided us a reference for many of the specific rules of the inspecting device that is tested here. Varian (1990) reports the case of the Grameen Bank in Bangladesh, where potential borrowers mutually monitor each others projects to ensure the success of the financed enterprise.

The aim of the paper is twofold. On one side, we want to replicate the Walker, Gardner, and Ostrom (WGO) experiments (1990) in a common-pool resource environment without sanctions. They find that in the lab the tragedy of the commons is more severe than was expected. This result is in sharp contrast with the excess of cooperation in public good provision experiments (Ledyard, 1995). On the other side, we want to study the influence of a self-financed monitoring and sanctioning institution on the efficiency of the resource use. It turns out that an inspecting mechanism greatly improves the group performance. The classical model with homogeneous, self-interested agents largely fails to account for the experimental results. Individuals are heterogeneous and this fact affects also the aggregate outcome. We present a model with heterogeneous, other-regarding agents that captures in a parsimonious way the main features of the data.

The reader is introduced to the topic in sections 2, 3, and 4. The theoretical predictions of the two models are outlined in sections 5 and 6. We then present the result of the experiments without sanctions (Section 7), with weak sanctions (8), and with strong sanctions (9). The conclusions follow (10).

2. DESCRIPTION OF THE PROBLEM

We have a group of N agents who can either invest their endowments E in a constant return activity or in the appropriation of a common-pool resource. The payoff of a generic agent i is the following:

$$\pi_i = v\cdot(E - x_i) + \frac{x_i}{X}\cdot f(X) \qquad \text{where } X = \textstyle\sum_{i=1,\dots,N} x_i \qquad (1)$$

In the investment decision, the constant-return activity plays the role of the opportunity cost that the agent faces when he shifts his effort to the appropriation of the common-pool resource. For instance, a peasant can either spend his working year cultivating his private arable land or raising cattle on the common pasture, or allocate part of his time to both activities.

The total flow of revenues from the common resource has been modeled with a technology $f(X)$. The group revenues $f(X)$ increase in $X \in [0, X'']$ and decrease in $X \in [X'', +\infty)$. What matters to the group are the revenues once the opportunity cost of the alternative activity is subtracted, $W = f(X) - vX$. The size of W – that we will call 'pie' – is maximum at some level of group investment below X'', but this level will not be the outcome if agents are self-interested because there is an incentive to deviate. The payoff of agent i π_i can be depicted as his slice of the pie, which depends both from the size of the pie W and from the share of it he can obtain, which is proportional to his individual investment relative to the group ($\frac{x_i}{X}$). If the group investment is at the socially optimal level, an agent profits from increasing his investment x_i, because he appropriates a larger share of the pie. As a consequence of this action, the pie shrinks in size causing a negative externality for the whole group (see Figure 1). A fisherman who over-fishes today, for instance, will reduce the tomorrow's catch for him and for his peers.

In conclusion, the unstructured interaction of N self-interested agents will lead to an excessive use of the common resource. This sub-optimal outcome is known as 'tragedy of the commons' (Hardin, 1968)[2].

This framework models important real world situations such as the exploitation of renewable resource, the use of a shared central computer, and the Cournot competition in an oligopoly[3].

< *Figure 1 about here* >

The revenues from renewable resources such as fisheries or forests are modeled in the literature with a technology $f(X) = a \cdot X - b \cdot X^2$ (Gordon, 1954; Clark, 1976), where X is the total harvesting effort on the resource and $a, b > 0$. If a group of N people can access the resource, then the appropriator i will get $\frac{x_i}{X} \cdot f(X)$ units of harvest (fish or timber, in the examples).

Another example is a group of users connected in a local net that share the CPU time of a central computer. Everybody sends to the central machine some tasks to be performed $x_1 \ldots x_N$. The amount of work done by this computer $f(X)$ will increase as the number of tasks submitted increases. After a given point, however, there will be congestion and the machine will waste a lot of time simply switching back and forth from one user to another and will actually decrease the amount of net time spent in solving users' tasks. From the point of view of the user, the amount of work that they will get done depends from the speed of the machine $f(X)$ as well as from the share of CPU time x_i/X they get.

3. LITERATURE REVIEW

The main purpose of this paper is to analyze the effects of a self-managed monitoring and sanctioning mechanism in a common-pool resource environment. The parameters of the no sanction common pool resource game have been taken – with some adjustments – from the experimental study of WGO. Their results from a sanction-free environment can be summarized in two points[4]:

---

[2] As a terminology clarification, notice that a free rider *over*-invests in appropriation efforts in a common-pool resource environment while he *under*-contributes in a public good provision game.

[3] The payoff π_i could be seen as the profits of an oligopolistic firm i that decides the quantity of goods x_i to produce and sell on the market. If we assume that the market demand is a linear function $D = a - b \cdot X$ and the technology has constant return to scale, i.e. the cost to produce a unit of the good is always v ($e = 0$), then the total revenues in the market are $f(X) = a \cdot X - b \cdot X^2$ and the share of the revenues that goes to firm i is proportional to its market share x_i/X.

[4] Conclusion 2 is an interpretation given in Ostrom et al, 1994, p.117.In their study there is a third conclusion: the total group investment exhibits a pulsing pattern across periods. Investment levels decayed and then rebounded repeatedly in the same experiment. When the idea of 'pulsing patterns' is operationalized into precise indexes, we find systematically lower values in our no sanction experiments compared to theirs. Having said that, we don't have any benchmark to state that our results did or did not exhibit pulsing patterns.

(1) the agents heavily over-use the resource at levels that go beyond what a pure free-riding behavior would suggest;
(2) The pattern of individual appropriation levels does not stabilize at the one-shot Nash equilibrium

The excessive appropriation efforts were particularly impressive when the maximum potential effort of the agents was very high in relation to the Nash equilibrium appropriation level (about three times as much)[5]. This result is puzzling because in voluntary contributions public goods games agents exhibit cooperation levels in excess of the Nash equilibrium prediction (Ledyard, 1995) while WGO reports that in common-pool resource games the cooperation level is below it.

In this paper we introduce some changes on the descriptive and methodological variables of the original study in order to test if we can push the behavior closer to what is predicted by the symmetric Nash equilibrium. In our no sanction design the monetary incentives have been substantially increased compared to the original study, by reducing the minimum safe earning and raising four times the conversion rate between laboratory currency and dollars. We have rewritten both the instructions and the software. Other parameter adjustments have been made to facilitate the understanding of the experiment by the subjects, such as re-scaling the action space, but without compromising the possibility of comparing the results with the original experiments. For a complete list of changes, please see the Appendix.

We replicated the qualitative results described above.

A self-managed monitoring and sanctioning institution was then added to the design to test its efficacy and efficiency in promoting cooperation among the agents. Many field studies suggest that a formal sanctioning system was often adopted by the group to keep the excessive use of a common resource under control (Ostrom, 1990). While the group outcome is common knowledge, monitoring enables the agents to observe the individual actions of other agents. Sanctioning enables an agent to reduce the payoff of other group members. An inspection involves both a monitoring action and a sanctioning action. The self-administered sanctioning devices already tested in the experimental literature are systems of 'social sanctions' that are defined by the following two features. Usually, the action of punishing is a loss for both the person who inflicts it and for the one targeted by it. Sanctions constitute a *complete* deadweight loss for the group. In addition, the more an agent is willing to pay for sanctioning the harsher the punishment inflicted. An example of social sanctions is the damage that a member of a village inflicts overnight to the nets of a fellow fisherman.

A social sanction system is characterized by a reward parameter ϑ, that measures the amount of punishment inflicted (s) relative to the cost to request a sanction (c): $\vartheta \equiv s/c$, where $c, s > 0$. We have examined three studies of self-administered social sanctions (Ostrom at al, 1994; Moir, 1998; Fehr and Gatcher, 1999) and have found the following common results[6]:

(1) Agents also request sanctions when it is costly for them to do so ($c > 0$).
(2) The lower the reward of the sanction (ϑ), the lower the number of requests to sanction.
(3) When agents can inflict sanctions, the group gross efficiency improves considerably.

---

[5] Walker, Gardner, and Ostrom (1990) test two designs with different ratios of Nash equilibrium over the endowment. In the high endowment case (the one that we have replicated here) the ratio is 0.32 and in the other case is 0.64.

[6] Other studies have been not included in the review either because they have an external sanctioning authority (Beckenkamp and Ostmann, 1999, Cardenas et al, 1999) or because the experimental design is for other reasons too different from ours (Yamagishi, 1988; McCusker and Carnevale, 1995). Experiments with external sanctioning authorities reports a less succesful story in raising group efficiency. The authors attributes at least part of the modest performances to the psychological distance between the users and the sanctioning authority.

The very fact that some agents request a social sanction (Point 1) contradicts a model of purely self - interested agents.[7] Agents are however responsive to the effects of sanctions on other people's earnings (Point 2). Since free riding actions seem to be more often targeted by sanctions than cooperative actions, the group outcome improves (Point 3) [8]. When the sanctioning fee c and the punishment s are subtracted from the gross efficiency figure, we obtain net efficiency. The effect of social sanctions on net efficiency is not yet clear and it probably depends upon the parameters of the sanctioning technology. Ostrom et al (1994) report a net loss while Moir (1998) and Fehr and Gratchen (1999) find a slight net gain in the group performance. In this study we find support especially for results (1) and (3) mentioned above.

Our design implements a 'legal sanction' system that presents two fundamental differences with a ' social sanction' system:

- The punishment (s) is a transfer of money from the targeted agent to the inspector. Hence inspecting is profitable when the fine is higher than the sanctioning fee (c), which remains a deadweight loss for the group.
- The punishment needs to fit the crime and does not depend upon the will of the inspector. In the specific case, the agent inspected pays a fine only if he invested more than a publicly known level and the fine is proportional to his excessive investment.

There are two additional features of our design that are different from other studies. An agent can be convicted only once for the same violation. In other words, there is no cumulating of sanctions that are requested by different agents on the same action. An inspection involves at the same time information discovery as well as punishment. In our design the individual actions are unknown before requesting the inspection: monitoring and sanctioning occur at the same time. Table 1 provides a systematic comparison of the four designs.

< ***Table 1 about here*** >

The features of the inspection device were crafted after a field study about the organization of rural communities in a mountain region of Italy, where a sanctioning mechanism was in place to limit overexploitation of the village forests and pastures (Casari, 1997). These institutions survived for more than six centuries. Both the conditions under which a sanction could be inflicted and the amount of the fine were specified in advance in written documents. The village court would sentence people who used the common resource above an established limit to pay a fine proportional to the severity of the damage inflicted to the community. A share of such fine usually went to the prosecutor. Any villager could report a violation but he usually incurred in a monitoring effort to discover the free rider and some costs to bring him to court. Other similar case studies are described in Ostrom (1990).

---

[7] In a social sanction system the best action of a self-interested agent is not to request sanctions at all. Experimentally, the use of sanctions might simply be the result of a trembling hand behavior or plain confusion. As it will become clear, in a legal sanction system the equilibrium can be moved away from the border of the action space.

[8] In Fehr and Gatcher (1999) targeting free riders is more rewarding than targeting cooperators. In Ostrom et al (1994) and in Moir(1998) this sanctioning behavior seems dictated by norms of fairness since there is no difference in reward across players. In Fehr and Gatcher (1999) the statement that the reward ϑ from punishment is higher for cheaters than for cooperators is not explicitly made. From simple calculation, it is easy to see that $\Delta\vartheta=1.2$: if everybody else fully cooperates, then $\vartheta=3.2$ when the agent fully cooperates and $\vartheta=4.4$ when the agent fully defects. When everybody else fully defects, then $\vartheta=0.8$ when the agent fully cooperates and $\vartheta=2$ when he fully defects.

In the simplified version for the laboratory, the inspection fee is the equivalent of these procedural costs and the entire fine is transferred to the prosecutor. In our experiment, the verdict of the court is always supposed to be correct and the identity of the prosecutors is not revealed. Moreover, there is no communication face to face.

4. EXPERIMENTAL PROCEDURES

There are 8 players, each of whom faces an investment decision in a single market that yields variable returns. A simultaneous move game is played where each player can choose an investment level from zero to a maximum level. A constant unitary cost is charged to the players for every token invested. The *gross group return* $f(X)$ from the market depends in a non-linear fashion on the sum of the investments X of all the players and is first increasing and than decreasing in X. The individual return consists of a share of the total group return corresponding to the fraction of the individual investment x_i on the group investment X.

The subjects were 56 students recruited from the campus of the California Institute of Technology. The experiments were run on networked personal computers using dedicated software for Netscape. Each subject was paid privately in cash immediately following the experiment. An experiment lasted from 1 hour to 2 hours and 20 minutes including the preliminaries (instructions, questions and answers, quiz, and practice rounds). Individual earnings ranged from $5.80 to $53.10. All subjects were seated at terminals, separated by partitions, and assigned identification numbers. Instructions were read aloud to everyone. The players could invest " tokens" and the account of payoffs was kept in francs (an artificial laboratory currency with a publicly known dollar-exchange rate) and in dollars.

An experiment consisted of a number of periods and each period consisted of just one step in the no sanction treatment and of two steps in the sanctioning treatments. During step one, the computer screen prompted a request for a number of tokens that the player wished to invest in the market. A player could digit any real number between 0 and 50. After everybody completed the input, the software displayed the group outcome (total group investment and gross group return). At this point, in the no sanction treatment players could also see their individual period payoff (your share of gross, cost of tokens, period payoff), while in the sanctioning treatments this part was postponed until the end of step two. Step two gave a chance to inspect other players. By clicking on a box next to the player identification number, a subject could ask to uncover the investment level of any number of players from 0 to 7. A fee was charged for every inspection and the eventual fines collected were credited to the inspectors. After everybody had taken this decision, the period results were displayed.

The period payoff was computed and explained into its three components: result of investment decisions, result of inspections asked, and notices of the eventual charge for an inspection targeting the player. In case more than one player asked to inspect the same person, a random device would pick only one inspector and cancel the requests of the others. At this point, the investment level of the players inspected during that period becomes public information. The software was designed in a way that a history record of the decisions always appeared on the computer screen. Players could see their past individual investments, the past total group investments, the past gross group returns, and the past uncovered investments of players inspected, and their individual cumulative payoff.

To ensure that the rules of the game were well understood we used the following procedure. First, the rules were publicly explained in detail and with examples. Second, a quiz was given. All the correct answers were read aloud after completion of the quiz and the ones where mistakes were noticed in the answers were further explained. Third, two practice periods were played, to help the subjects familiarize themselves with the rules of the experiment and with the software. After the two practice rounds, a number of periods from 27 to 33 were run. Players were not told the number of rounds they were going to run. At the end of the third -before-the-last period, an announcement was made that the experiment was going to end in two periods. After the experiment was over, a questionnaire was submitted to the subjects asking for the strategy they followed.

A total of ten experiments were run, grouped into three different treatments. Four experiments were run with no inspection device (No Sanction) and six with an inspection device. Of this last group, one set of sanctioning parameters were used for four experiments (Weak Sanction) and a different set for the other two (Strong Sanction). Within each treatment, half of the experiments were conducted with inexperienced subjects and the other half with experienced subjects.

5. CLASSICAL MODEL

In this section we compute the Nash equilibrium of the game under the three experimental treatments – no sanctions, weak sanctions, strong sanctions - using a standard model of homogeneous, self-interested agents. We will refer to it as the classical model. We assume that the agents are risk neutral and that the preferences of all the agents are common knowledge.

5.1. NO-SANCTION DESIGN

The payoff function of a self-interested agent i in the no-sanction design is:

$$\pi_i = v\cdot(e - x_i) + \frac{x_i}{X}\cdot(aX - bX^2) \qquad \text{where } X=\Sigma_i x_i \qquad i=1,\ldots, N \qquad (1')$$

The best response function is a linear function of the investment level of everybody else, $x_{-i} = \sum_{j\neq i}^{N} x_j$:

$$x_i^* = \frac{a-v}{2b} - \frac{1}{2}x_{-i} \qquad \text{where } x_i^* \text{ is bounded to be in } [0, 50] \qquad (2)$$

< *Figure 2 about here* >

> **Proposition 1A**. (*RESOURCE USE EFFICIENCY WITHOUT SANCTIONS*)
> Without a sanctioning institution the Nash equilibrium with homogeneous, self -interested agents [X=128] has an efficiency of 39.5% of the optimal level.
> All the agents will use the resource at an identical rate of 16 tokens.

When agents are identical, the Nash equilibrium of the game is $x_i = \frac{a-v}{(N+1)b} \forall i = 1,\ldots, N$ that - given the parameter values N=8, a=23/2, v=2.5, b=1/16 - corresponds to an individual investment of x_i=16 and to a total group investment of X=128. The socially optimal outcome is at X=72 could be obtained if all the eight agents in the group invest 9 tokens [9]. In order to make a welfare comparison across experiments, we define an efficiency index that assumes values of 100 and below. The maximum level of 100 is reached when the net return is equal to the socially optimal outcome (π=324, where the endowment is set to zero, E=0). The index is the percentage of the net group return compared to that maximum value. The Nash equilibrium has an efficiency of 39.5%. The

---

[9] In the parameter setting that has been chosen the individual and group earnings are very sensitive to the presence in the group of even a single "crazy" player. If – for instance - seven people stick to their Nash equilibrium share of investment (x_i=16) and the eighth agent invests the maximum amount (x_i=50), the net group return declines from $+8.32 to $-4.09.

value is 0% at the open access level. At the open access investment level of X=144 (x_i=18) the net group return is equal to just the period endowment, which implies a complete destruction of the potentially positive profits that the group could have made out of the common-pool resource. This outcome is the equilibrium level at the limit when the number of appropriators goes to infinity and it is the analogous to a perfect competition in a market setting (Figure 1 depicts the three reference points).

5.2. WEAK SANCTION DESIGN

In the sanction treatments each period has an investment and an inspection phases. Once the total group investment is revealed, agent i might ask to inspect any other agent j (I_{ij}=1 when an inspection is requested, I_{ij}=0 otherwise, i≠j). A fee k is paid for every inspection carried out and a transfer s_j from agent j to agent i is made. The transfer s_j is proportional to the investment of agent j in excess of a legal threshold of λ tokens. The total revenues from inspections for agent i are $r_i = \sum_{j=1}^{N} I_{ij} r_{ij}$, where:

$$r_{ij} = s_j - k \qquad \text{and} \qquad s_j = \begin{cases} 0 & , x_j \le \lambda \\ h\cdot(x_j - \lambda) & , x_j > \lambda \end{cases}$$

The parameter h is the unitary fine for each extra token invested and measures the stiffness of the punishment. Sanctions modify the incentives for investment because they threat to increase the cost of tokens invested above a given limit λ. In Figure 2, sanctions induce a downward shift in the best response function of a targeted agent. The degree of the shift depends on the perceived probability p_i that an agent has of being inspected. If such probability is zero for all the agents, the game is identical to the no sanction one but if it is strictly positive for some agent, her best response is to invest less than in the corresponding no inspection case:

$$x_i{*} = \frac{a - v - h\,p_i}{2b} - \frac{1}{2}x_{-i} \qquad , \quad if \quad x_i \ge \lambda \qquad (4)$$

> **Proposition 2A**. (*RESOURCE USE EFFICIENCY WITH WEAK SANCTIONS*)
> When agents are homogeneous and self-interested, the introduction of weak sanctions will not change the classical Nash equilibrium level stated in proposition 1A.
>
> **Proposition 3A.** (*INSPECTION BEHAVIOR WITH WEAK SANCTIONS*)
> When weak sanctions are introduced, no inspection will be requested by homogeneous, self-interested agents.

When weak sanctions are in place (k=7, λ=9, h=1), the symmetric Nash equilibrium in the investment phase and in the inspection phase ranges in $X^* \in$ [113.7, 128] - or equivalently $x_i^* \in$[14.2, 16] - depending on the value of p. There is an assumption of p_i=p ∀i. As the probability of being inspected p increases, the total group investment in equilibrium X^* decreases.

At this point, the question is what determines the value of p_j, the probability of agent j of being inspected. An agent j will be inspected if it is profitable for another agent i to do so, or if $r_{ij} > 0$, which occurs when the investment level of agent j is sufficiently large, i.e. $x_j > \underline{x} \equiv \frac{k}{h} + \lambda = 16$. The sanction can be inflicted when the investment is above a "legal" threshold of λ=9 but it is not profitable to actually exercise that right until the over-investment is sufficiently large, namely x_i>16 because the inspector has to pay a fixed fee k=7. With symmetric agents, no inspection will be profitable until $X \leq 128$ and when $X > 128$ the inspection of each one of the agents will be profitable. The number of requests to inspect agents will be zero before $X=128$ and jump to $N(N-1)$ (i.e. 56 for $N=8$) after that point.

An equilibrium investment level X* and probability p* are jointly determined by the two phases of the game. The symmetric Nash equilibrium is $X^*=128$ and $p^*=0$. To sum up, if all agents are self-interested and that fact is common knowledge, then the total group investment will not change from the no sanction design level and there will be no inspection[10].

5.3. STRONG SANCTION DESIGN

The weak sanction design does not alter the equilibrium outcome of the no sanction design. The inspecting device simply puts stronger incentives to discourage over-investment, which is not in the agents' self interest anyway. The strong sanction design has the explicit purpose to move the equilibrium away from the inefficient equilibrium of the no sanction treatment to an efficient equilibrium.

> **Proposition 4A** (*RESOURCE USE EFFICIENCY WITH STRONG SANCTIONS*)
> When agents are homogeneous and self-interested, the introduction of strong sanctions will move the Nash equilibrium very close to the socially optimal level [above 99% efficiency, X=71.1].
>
> **Proposition 5A**. (*INSPECTION BEHAVIOR WITH STRONG SANCTIONS*)
> When strong sanctions are introduced, all agents will inspect everybody.

There are many ways to modify the inspection parameters k, h, λ in order to move the equilibrium to the socially optimal point of $X=72$. The technological parameter k (inspection fee) that represents ideally the degree of difficulty in observing other people's actions was not changed. Instead the legal parameters h and λ were adjusted, by inflicting stronger punishments for violations of tougher individual quotas ($h=4$, $\lambda=7$, $k=7$).

For the new set of parameters, the equilibrium group investment ranges in $X^* \in [71.1, 128]$ - or equivalently $x_i^* \in [8.9, 16]$ - depending on the value of p. Inspecting an agent is profitable when $x_i > 8.75$. If agents are symmetric, no inspection will be profitable unless $X > 70$. In the new symmetric equilibrium (X*, p*) is (71.1, 1) all the agents are inspected and the group efficiency is at 99.97%. The total group investment X* is slightly different from the social optimal value because the preference was to keep the parameter values integer numbers in order to facilitate computation by the agents. The difference in terms of efficiency is, however, negligible.

6. A MODEL WITH HETEROGENEOUS, OTHER-REGARDING AGENTS

In this section we outline a simple model where an agent's payoff depends not only on her personal earnings but also on the earnings of the other people in the group and then compute the Nash equilibrium of the game. This class of models appears both in some of the early texts of the classical

---

[10] The pair $X^*=128$, $p^*=0$ is still the equilibrium when the agents have a "trembling hand" in their investment decisions. In equilibrium the agents are indifferent between inspecting and not inspecting and hence it is not a robust equilibrium when the " trembling hand" is in the inspecting decisions. If a player inspects "by accident" and this kind of events is common knowledge the equilibrium will be below $X=128$. We believe that neither our basic results nor our conclusions are affected by this point.

and marginalist schools (Smith, 1759; Edgeworth, 1881) and in more recent experimental works (Krebs, 1970; Rabin, 1993; Ito et al, 1995; Fehr and Schmidt, in press; Bolton and Ockenfels, 1999; Levine, 1997; Chan et al, 1997). The recent, growing interest in these models stems from a realization that " pure self interest is clearly not a fully adequate description of human motivation. Realism suggests that economists should move away from the presumption that people are motivated *solely* by self interest" (Rabin, 1996). Rabin (1996) cites an extensive body of experimental research where people exhibit pattern of not self-interested behavior.

The specific shape of other-regarding preferences presented in this section intends to capture in a parsimonious way a component that is believed to motivate human behavior.

MODEL WITH HETEROGENEOUS, OTHER-REGARDING AGENTS

$$U_i(\pi_i, \Pi_{-i}) = \pi_i + \gamma_i\, \Pi_{-i} \qquad (3) \quad \begin{cases} \gamma_i > 0 & \text{Altruistic agent} \qquad \Pi_{-i} = \Sigma_{j\neq i}\pi_j \\ \gamma_i = 0 & \text{Self-interested agent} \qquad \gamma_i \in [-1,+1] \\ \gamma_i < 0 & \text{Spiteful agent} \end{cases}$$

Self-interest is a special case of the model ($\gamma_i = 0$). In general, agent i is willing to give up \$1 of personal earnings (π_i) in order to see the other people's earnings (Π_{-i}) changed by 1/ γ_i dollars. A positive value in the other-regarding parameter γ_i denotes an altruistic attitude toward the group, while a negative value denotes a spiteful attitude. *A spiteful agent will find enjoyment in decreasing the earnings of others* and therefore she is willing to use some of her personal earnings in order for that to happen. The degree of altruism or spite is bounded in a way that nobody will choose to pay more than \$1 to modify the group earnings by less than \$1. Although not crucial for the conclusions, we think that this assumption of $\gamma_i \in [-1, +1]$ is reasonable. The definition of spite we have given is similar to the one adopted by Levine (1997) but different to the concept of envy suggested by Mui (1995). The model (3) does not incorporate any reciprocity nor equity nor fairness considerations [11].

We will study this model under the assumption of group heterogeneity that we consider agents with different degrees of concern about others' earnings γ_i and, in particular, groups where at least two agents i, k are different, $\gamma_k \neq \gamma_i$. As in the classical model, we assume that the agents are risk neutral and that the preferences of all the agents are common knowledge . We will refer to (3) as the model with heterogeneous, other-regarding agents.

The remaining of this section is devoted to the computation of the Nash equilibrium in the three designs[12]. The payoff function for agent *i* in the no sanction design is now:

$$\pi_i = v\cdot(e - x_i) + \frac{x_i}{X}\cdot(aX - bX^2) + \gamma_i\,\left[v\cdot((N-1)e - X_{-i}) + \frac{X_{-i}}{X}\cdot(aX - bX^2)\right] \qquad (1'')$$

where $X=\Sigma_i\, x_i$, $X_{-i}=\Sigma_{j\neq i}\, x_j$. The best response function is:

$$x_i{*} = \frac{a - v}{2b} - \frac{(1+\gamma_i)}{2}x_{-i}$$, where $x_i{*}$ is bounded to be in [0, 50]. (2'')

---

[11] This is not a denial in the role of reciprocity, equity or fairness considerations. For the sake of simplicity, we included in the model the component that we believe is most important.

[12] To simplify computation, we assume that the vector of other-regarding parameters γ is such that the *individual* response function is within the interval $x_i \in [0,50]$. This assumption might further restrict the range of γ to a subset of the [-1, +1] interval.

Proposition 1B. (*RESOURCE USE EFFICIENCY WITHOUT SANCTIONS*)
Without a sanctioning institution the Nash equilibrium with heterogeneous, ot her-regarding agents depends upon the preference structure of the agents. For instance:

- When *all* agents are altruistic, group efficiency will be better than the classical Nash (>39.5%).
- When *all* agents are spiteful, group efficiency will be worse tha n the classical Nash.

In general, individual agents will use the resource at a different rates.

In the no sanction environment (N=8, a=23/2, v=2.5, b=1/16), The *symmetric* Nash equilibria of the game range in $X^* \in [72, 400]$ that corresponds to an efficiency interval [–321%, 100%] depending on the value of the other-regarding parameter vector γ. For illustrative purposes, if all agents are homogeneous and altruistic with $\gamma_i = 1/7$ then the Nash equilibrium is $X^*=115.2$, while if they are all identically spiteful with $\gamma_i = -1/7$, then $X^*=144$ [13]. With heterogeneous preferences the outcome is still in the range indicated above but there is no longer a one -to-one correspondence between a group outcome X* and a unique vector of agent preferences. An interesting case is when the preferences in the group are symmetrically heterogeneous or, in other words, for every altruistic agent *i* with $\gamma_i > 0$ there is a spiteful agent *k* with $\gamma_k = -\gamma_i$. The Nash equilibrium with symmetrically heterogeneous agents is in general more efficie nt than the classical Nash equilibrium (39.5%).
With heterogeneous preferences also the individual appropriation levels are heterogeneous. In particular, the lower the other-regarding parameter γ_i, the higher the individual appropriation x_i is: spiteful agents use the resource more than self-interested agents and self-interested agents use it more than altruistic ones.

Proposition 2B. (*RESOURCE USE EFFICIENCY WITH WEAK SANCTIONS*)
When agents are heterogeneous and other -regarding, the introduction of weak sanctions will improve upon the Nash equilibrium level without sanctions stated in proposition 1B.
If two or more agents are not altruistic, the improvement will be strict and inequality in the individual use of the resource will decrease relative to the no sanction design. (sufficient condition)

Proposition 3B. (*INSPECTION BEHAVIOR WITH WEAK SANCTIONS*)
When weak sanctions are introduced,

- there will be a positive number of inspections if two or more agents are not altruistic (sufficient condition)
- spiteful agents will be more aggressive inspectors than altruistic agents and will also purposively request non-profitable inspections.

When agents are heterogeneous and other-regarding, they use the resource at different rates (proposition 1B) and in pa rticular at least one non-altruistic agent uses the resource above $x_i=16$. If there are two or more self-interested or spiteful agents, that action will be inspected. Since an

---

[13] With all self-interested agents $\gamma_i = 0$ the outcome is the same as in the classical model, namely $X^*=128$. If all the agents are homogeneous, the group appropriation level with other-regarding agents is $X^*=72$ with completly altruistic agents ($\gamma_i = 1$) and is $X^*=400$ with completely spiteful agents ($\gamma_i = -1$).
The game has of course also asymmetric Nash equilibria, which depend upon the individual preferences within the group. For instance, when half of the agents are altruistic $\gamma_i = 1/7$ and half are spiteful $\gamma_i = -1/7$ (symmetrically heterogeneous preferences) the group outcome is $X^*=126$ and the individual investment levels will be $x_i = 0$ and $x_i = 31.5$ respectively.

inspection will turn out in an extra cost proportional to the over-use of the resource, in a sanction environment at least one agent has an incentive to decrease her resource use level. As a consequence, the group appropriation rate decreases and the welfare improves compare to the no-sanction environment.

Under the threat of sanctions, the best response function of an agent i with other-regarding preferences is:

$$x_i^* = \frac{a - v - p(1+\gamma_i)h}{2b} - \frac{(1+\gamma_i)}{2} x_{-i} \quad \text{, if } x_j > \lambda \qquad (4')$$

Spiteful agents are particularly sensitive to the threat of sanctions because the money that is taken away from them is given to somebody else while altruistic agents care less about sanctions precisely because of this fact. Sanctions induce spiteful agents to lower their investment level proportionally more than altruistic agents. In particular in the weak sanction design ($h=1$, $\lambda=9$, $k=7$), the inequalities in investment levels within the group will be reduced, although spiteful agents will still use the resource more than altruistic agents will.[14]

Agent i will inspect agent j when $x_j > \frac{k}{h\,(1-\gamma_i)} + \lambda$ and, of course, $x_j > \lambda$[15]. Inspection decisions are affected by the value of the other-regarding parameter γ_i. In particular a spiteful agent *i* inspects for lower values of x_j than an altruistic agent does. The reason is that she finds enjoyment not only from the cash flow of the fine but also from decreasing the income of some other agent. On the other hand, an altruistic agent is concerned about the social loss constituted by the inspection fee *k* and does not consider all the money of the fine s_j as a gain since it has been subtracted from somebody else she cares about. For instance, in the weak sanction treatment, a moderately spiteful agent with $\gamma_i = -1/7$ will request an inspection for any $x_j > 15.1$ (and lose money if $x_j < 16$) compared to a self-interested agent who would do it only when $x_j > 16$. A moderately altruistic agent with $\gamma_i = 1/7$ would inspect only agents with $x_j > 17.2$. A completely altruistic agent ($\gamma_i = 1$) will never inspect, while a complete spiteful one inspects when $x_j > 12.5$. An implication of this examination is that when facing the same investment pattern, high investors will be more aggressive inspectors than low investors will.[16]

Proposition 4B. (*RESOURCE USE EFFICIENCY WITH STRONG SANCTIONS*)
When agents are heterogeneous and other-regarding, the introduction of strong sanctions will shift the Nash equilibrium at an efficiency level above 98% of the socially optimal level under some regularity conditions on preferences ($X \in [64, 72]$).

Proposition 5B. (*INSPECTION BEHAVIOR WITH STRONG SANCTIONS*)

(i) When strong sanctions are introduced, all agents will be inspected under some regularity conditions on preferences

(ii) Spiteful agents will be more aggressive inspectors than altruistic agents.

---

[14] Individual earning inequalities from resource use (excluding revenues from the inspection activity) will also decrease in the periods where the profits of the group are non-negative, $f(X) - vX = 0$ (i.e. $X \leq 144$).

[15] The payoff function of agent i when he can inflict a sanction on agent j is $U^{\wedge}_i(\pi_i, \Pi_{-i}) = U(\pi_i, \Pi_{-i}) + (s_j - k) - \gamma_i s_j$. The decision is to inspect when $U^{\wedge} > U$, or $(s_j - k) > \gamma_i s_j$,

[16] Suppose that an inspection is unprofitable. The cost for the inspector is $c = k - s$ and the damage inflicted is s. So in our other-regarding model agent j will request the inspection if $\vartheta < -\gamma_i$, where $\vartheta = s/c$.

In a strong sanction environment (k=7, λ=7, h=4), the symmetric Nash equilibrium is in the range $X^*\in$ [64, 72] - or equivalently $x_i^*\in$[8, 9] – when the probability of being inspected p is set equal to 1. If all agents are completely altruistic (γ_i =1) the outcome will be socially optimal (X=72, 100% efficiency) while if all agents are completely spiteful (γ_i =-1) the outcome will be 98.77% efficient (X=64).

In equilibrium the probability of agent j being inspected p_j, is actually equal to one, under some regularity conditions on preferences. A completely altruistic agent (γ_i =1) will never inspect while a complete spiteful one inspects when x_j>7.9. A sufficient condition on group preferences for all actions to be inspected is that there are two or more non-altruistic agents and that the most spiteful one is not too far apart from the next. More formally, when agents are ranked low to high other regarding parameters $\gamma_{(1)}, \gamma_{(2)},\ldots, \gamma_{(8)},$ then $\gamma_{(1)}, \gamma_{(2)}\leq 0$ and $|\gamma_{(1)} - \gamma_{(2)}| < 0.25$.[17]

In the weak sanction treatment spiteful agents still invest more in absolute terms than altruistic agents while with strong sanctions, the situation is reversed because of the stiffness of the sanctions: the higher investors are relatively more altruistic than the lower investors. According to the heterogeneous, other-regarding agent model, the lowest investors will be more aggressive inspectors than the highest investors will.

7. RESULTS OF NO SANCTION EXPERIMENTS

The experimental results are compared with the predictions of the classical model (1A) and of the heterogeneous, other-regarding agent model (1B). The data do not support the classical model predictions in the no sanction environment.

> **Result 1A**. Without a sanctioning institution the resource is overused relative to the Nash equilibrium with homogeneous, self-interested agents ('classical' Nash equilibrium). People cooperate less than expected according to that model and are worse off than the model predicts.
> (i) Actual resource use is greater than the classical Nash equilibrium.
> (ii) The phenomenon is not explained by learning or experience.

Support:

(i) In terms of efficiency the groups scored 28.4% of the maximum possible net return, a value that is in-between the classical Nash equilibrium level of 39.5% and the open access level of 0%. The overall average of the group investment for the four experiments was 131.3, which is statistically different from both the above reference values at a 0.01 level [18] (see Table 2 for details). The group investment varied considerably across periods, ranging from a minimum of 85.5 to a maximum of 167 tokens.

(ii) Learning or experience effects do not alter the main conclusion that the group investment is persistently above the one-shot classical Nash equilibrium level. There is no support in the data for the claim of temporary off-equilibrium outcomes due to learning or experience:

---

[17] The above condition if fully satisfied in all the no sanction experiments. A sufficient condition to have at least 87.5%(i.e. 7/8) of the actions inspected is that at least one agent is not strongly altruistic, γ_i <0.08.

[18] The symmetric Nash equilibrium value X=128 was never recorded in any of the 129 rounds in which the investment decisions were taken. The open access level is X=144.

- Experienced subjects do not perform better than inexperienced subjects do. Differences in efficiencies actually favor inexperienced subjects (25.2% versus 31.6%, Figure 3).
- A comparison between the first half, second half, and after announcement period averages[19] show no statistical differences at 0.01 significance level. As an overall average, the values are 131.37 in the first half, 131.39 in the second half, and 130.31 after the announcement (Figure 4).

The volatility of the group appropriation level decreases over time in three out of four experiments (see variance comparisons in Table 2) but it mostly reflects oscillations around the same average.

< ***Figure 3 about here*** >

We can further analyze the evolution of group appropriation over time. Agents know they are involved in a repeated game and they are not told the number of rounds they will go through. The probability that an additional round will be played decreases over time. If in the initial round agents can support a better resource use than the one-shot Nash equilibrium, such level will move progressively closer to the latter as the game unfolds. If a repeated game effect is present, the pattern in the total group investment should be

(a) A convergence to the one-shot Nash equilibrium from below, i.e. in the range $X \in [72, 128)$;
(b) An eventual jump to the one-shot Nash equilibrium level after the end-of-experiment announcement has been made.

The data show a weak support for (a) and no support for (b).

< ***Figure 4 about here*** >

The presence of a repeated game effect and of learning effects will be evaluated by the application on the data of a dynamical model described in Noussair et al (1995). The model assumes that total group investment may start from a different origin for each experiment, but the convergence is assumed to be a common asymptote in all four experiments. Formally the model is as follows:

$$X_{mt} = B_{11} D_1 \frac{1}{t} + B_{12} D_2 \frac{1}{t} + B_{13} D_3 \frac{1}{t} + B_{14} D_4 \frac{1}{t} + B_2 \frac{t-1}{t} + u_{mt}$$

Where m is the index of the experiment; D_k are dummy variables that take value 1 if m=k and value 0 otherwise; t is time measured in terms of experimental period number; X_{mt} is the total group investment in period t of experiment m. B_{1m} measures origin of the group investment convergence process, and B_2 is an asymptote. Data in Table 2 show the estimation of the model.

< ***Table 2 about here*** >

The asymptote for the no sanction experiments is 134.0, which is statistically different from the equilibrium level of 128 but not significantly different from the overall average group investment of 131.3 at a 0.05 level. This result confirms once more that the overuse of the resource persists and does not tend to die out. The convergence to the asymptote starts from below for all the experiments.

---

[19] In base no sanction experiments, the first half includes periods 1-15, second half 16-30 (or 16-31), and after announcement 31-32 (or 32-33). In sanction treatment, the first half includes periods 1-12, second half 13-25 , and after announcement 26-27.

Result 1B. A model of heterogeneous, other regarding agents explains the resource use data better than the classical model both at the aggregate and the individual levels. Furthermore, some clues exist in the data suggesting that heterogeneity is an appropriate modification to the classical model.
- Over investment and under investment are properties of individuals.
- Self-reported other-regarding preferences

Support: The observed level of group appropriation can be explained by the heterogeneous, other-regarding agent model given an appropriate pattern of group preferences that is biased toward spite[20].

Individual actions are very dispersed relative to the classical individual Nash equilibrium x_i=16 (a) and this variability is due to individual heterogeneity (b). Individual heterogeneity is not a consequence of confusion (c) but is consistent over time (d) and is due to other-regarding preferences (e), (f).

(a) The patterns of individual investment do not conform to the one-shot classical Nash equilibrium prediction. The actions within a 25% bandwidth around the prediction (i.e. in the interval [14, 18]) account for 15.7% of all the actions[21] and the rest are not symmetrically distributed around that value: about 61% are below and 23.3% are above. The mean is 16.42 and the standard deviation is 10.00.

(b) A brief look at the individual average investment levels makes clear that agents are heterogeneous and that only a few agents were accountable for a systematic over-investment (Figure 5). We can reject the hypothesis that the agent average investment is at the individual symmetric Nash equilibrium (xi=16) for 28 out of 32 agents at 0.05 level (see white bars in Figure 6). Within each experiment there are at least four different types of agents whom investment behavior is statistically different at 0.05 level. The presence of different types of individuals is a common finding in the experimental literature (Von Winden, Dijk, Sonnemans, 1998)

(c) There are reasons to believe that the differences in individual behavior are traits of the agents and are not due to confusion. The experimental design was not simple and a possible explanation of such behavior is that the " heavy investors" might have been confused players who did not properly understand the incentive structure of the game[22]. The evidence from the quiz completed by each player before the experiment does not show any support for this option. We have assigned a score to each quiz taken, which is 1 if all the answers are correct, 0.5 if some answers are not perfect but it is clear that the player overall understood the rules, and 0 if there are substantial and repeated mistakes. The four highest investors score an average of 0.92 against a general average of the 32 players of 0.89. In other words, the heavy investors seems – if something - better skilled than average.

(d) There is a remarkable consistency over time in the individual investment patterns, which indicates that the differences across agents are purposive rather than random. A rough measure

---

[20] Consider for example a group with three types of agents: two are moderately altruistic γ_i =1/21, four self-interested agents, and two quite spiteful ones γ_i =- ¼. The group investment is X*=132 with individual investments x_i of 6, 12, and 36 respectively.

[21] The actions exactly at x_i=16 are 26 (2.52%).

[22] If the heavy investors are confused players, however, it is unclear why we do not find them in the experiment with the weak sanction treatment (see Figure 5). Such experimental design is more complex than the no sanction design, although the threshold level for sanctioning gives a vague clue about the equilibrium level and the monetary incentive against high investment levels are higher.

of time consistency comes from the comparison of the agent average investment levels of the first and second half of the experiment. The greater the similarity between the two values, the stronger the consistency claim can be. In order to compare the correlation of the individual investment levels over time, we have run an ordinary least square regression on agent ranking. Each agent has been assigned her ranking position in terms of average investment within the experimental group. The first-half-of-the-experiment ranking has been regressed on the second-half-of-the-experiment ranking without a constant term. Rank correlation informs on the existence of any form of monotonic relation between the values and it is a better choice than absolute value correlation, which can capture only linear relations between the two variables. When using ranks, the coefficient is in the interval [-1, +1]. A negative value denotes a decreasing relation and a positive value denotes an increasing relation. The higher the absolute value, the stronger the monotonic relation is. A zero value means that there is no monotonic relation at all. Our best result will be a 1-value coefficient. When all the no sanction experiments are pooled together, the estimated coefficient is 0.936 (number of observations is 32, R-squared 0.88. See Table 2 for single experiment regressions). This test supports the view that over time agents are consistently heterogeneous.

(e) From here to make precise statements on the nature on unobservable preferences there is a jump. We find some help in the investment strategy notes that the participants left on their final questionnaire, which often mention other agents' earnings. Here is one: " My greed went to the extent of causing me not to want to fall behind having at least 1/8 of the market share because I didn't want others making more profit per period." One of the highest investors explicitly mentioned in the description of his strategy the goal of decreasing the earnings of the others: " [I] tried to keep a large portion of the token invested by forcing the others to adjust their investment so that the total would be profitable".

(f) The estimation of the other-regarding agent model on the experimental data leads to consistent results. Since the difference between the classical and other-regarding model is just in the slope of the best response function (cfr. (2) and (2")), the regressions assume a correct value for the intercept. Moreover, there is a conjecture that the agents expect the others to act in period t as they did in period t-1: $x_{i,t} = 72 - \frac{(1+\gamma_i)}{2} x_{-i,t-1} + \varepsilon_i$. All the 32 agent-specific estimated values of the slope fall into the allowed interval [-1,0] corresponding to an other-regarding parameter $\gamma_i \in [-1,+1]$. The γ_i estimates range from a minimum of -0.40 to a maximum of 0.08. About 37% of the agents have a parameter γ_i significantly different from zero at a 0.05 level and our model classifies them as either altruistic when γ_i is positive (2 agents) or spiteful when γ_i is negative (10 agents).

< Figure 5 about here >

The predominance of spiteful agents over the others can account for both the observed pattern of individual actions and for the overuse at the group level. A model relying on homogeneous, self-interested agents cannot explain either one of the two regularities [23].

---

[23] The experimental data might be explained by different specifications of other-regarding preference models. An example of specification different from the one hereby suggested is the *status seeker* model: $U_i(\pi_i, \Pi_{-i}) = \pi_i + \delta_i (\pi_i / \Pi_{-i})$, where $\delta_i>0$ when agents care about relative income and $\delta_i = 0$ for pure self-interested agents (see Ito, Saijo, and Une, 1995). The status seeker model does not have any room for "altruistic" actions; it has a very unnatural behavior for group appropriation levels above the open access level (X>144) because both π_i , $\Pi_{-i} < 0$. Regarding inspection decision,

8. RESULTS OF WEAK SANCTION EXPERIMENTS

This section describes the outcome of four experiments run under the weak sanction treatment and in particular it focuses on the inspection decisions (Result 3) and their effects on the appropriation system (Result 2).

> **Result 2.** With the introduction of weak sanctions,
>
> (i) Group efficiency improves substantially: the level of resource use moves from below the classical Nash Equilibrium to above the classical Nash equilibrium.
>
> (ii) Inequality in the individual use of the resource decreases relative to the no sanction design
>
> These results are not predicted by the classical model (Proposition 2A) but they are consistent with the heterogeneous, other regarding agent model (Proposition 2B).

Support:

(i) The efficiency level is considerably higher with weak sanctions than without sanctions and is well above the classical Nash equilibrium prediction (39.5%) for any specific index considered. The gross efficiency level has roughly doubled (28.4% without sanctions, 57.19% with weak sanctions). The change is minimal when we correct for the differences in length among experiments and consider the first 25 periods only (28.9% versus 56.21%).

Requesting sanctions is a costly activity and so a fair comparison needs to consider the cost of the inspection mechanism. In fact, inspection fees are a deadweight loss for the group and as such needs to be subtracted from the rent extracted under the sanction treatments. The amount of the fines is, instead, a plain transfer from an agent to another and is not a cost from the group standpoint. When the inspection fees (8.9%) are subtracted, the net rent is 48.3%, which is about twenty points above the no sanction level (cfr. tables 2 and 4). When the agents are experienced the efficiency improvement is even greater: the net rent of experienced subjects is on average 62.02% versus a 34.56% of inexperienced ones (see Figure 3).

The total group investment is substantially lower for sanction experiments than for no sanction ones. As an overall average, group investment drops from 131.3 to 115.5 tokens (statistically different at 0.01 level). The aggregate investment is statistically different from both the classical Nash equilibrium and the socially optimal level (0.01 level). The classical Nash equilibrium X=128 was recorded in the 1.85% of the rounds (2 out of 108). When considering the classical model, the overall group investment average is not statistically different from the one-probability inspection prediction (X=113.7) but it is one single-experiment average with experienced subjects (0.05 level). The group investment across periods ranged from a minimum of 87 to a maximum of 186.8, which is wider than the same range for no sanction experiments. Similar results come from the estimation of the Noussair convergence model explained in the Section 7. The ordinary least squared asymptote of X=114.8 is not significantly different from the one-probability inspection level (0.05 level) while it is from the zero-probability level (see Table 2).

---

agents never request inspections for $(\pi_i / \Pi_{-i}) < 1/\theta$, which implies that nobody inspects if they expect a loss. Moreover, poor agents request fewer inspections than rich ones (there is however a need to define wealth as current or cumulative earnings). A different behavior is predicted for different period endowment levels.

(ii) Agent inequality decreases compare to the no sanction treatment. The spread in average appropriation levels between the highest and the lowest investors greatly decreased with the introduction of weak sanctions and inequalities in appropriation levels measured using Gini coefficients fell dramatically (Table 2). As a direct consequence, income inequalities from appropriation decreased[24].

We need to assume that agent preferences in the no sanction experiments were not different than in the weak sanction experiments in order to interpret the results as supportive of the heterogeneous, other-regarding agent model.

> **Result 3.** With the introduction of weak sanctions, about half of the actions are inspected.
> (i) The number of inspections exceeds the number that is profitable. In fact twice the number of inspections take place than is ex-post profitable.
> (ii) The highest investors are more aggressive inspectors than the lowest investors
> These results are not predicted by the classical model (Proposition 3A) but they are consistent with the heterogeneous, other regarding agent model (Proposition 3B).

Support:

(i) The prediction of the classical model of no inspections is clearly incorrect, since 51.5% of the actions were inspected (Table 5). Some of the inspections turn out in a profit for the inspector while others in a loss (when x_i<16). The inspections with a negative balance were either mistakes due to the asymmetric information or purposive decisions of other-regarding agents. We will investigate both possibilities.

< ***Table 3 about here*** >

Given a group appropriation level, the agents know the number of over-users if they know the empirical density function of the agent types. In the first period, however, they don't know the identity of such agents.

The number of actual inspections compared to the number of potentially profitable ones is very high (about twice as many). This ratio does not decline as the agents are more familiar with the inspection device (1.86 inexperience versus 2.22 experienced subjects) and as the agents reveal their type during the experiment (2.02 first half of the experiment, 2.13 second half, 1.82 after announcement)[25].

< ***Table 4 about here*** >

We know that spiteful agents will purposely request some un-profitable inspections. Only inspections with a reward parameter ϑ below 1 can be safely classified as mistakes for that type

---

[24] Income from appropriation is defined as the gross revenues from appropriation minus the cost of tokens and it excludes the period endowment as well as costs and revenues from the inspection activity. There is a very strong linear correlation between average appropriation levels and average income from appropriation (0.97 for all experiments, 0.99 for no sanction experiments) and this fact explains the tight link between the Gini coefficients for the two variables. The standard deviation of average appropriation and income levels exhibit similar patterns.

[25] About 35.1% of the inspecting decisions were incorrect, either because a potentially profitable inspection has not been requested (type I error) or because a potentially non-profitable inspection has been requested (type II error). Such error rate is substantially lower for experienced agents (29.2% versus 41.0%). In spite of the excessive inspecting activity, the inspection balance is close to the zero level ([-1.5%, +1%] in terms of the maximum rent). Type II errors are higher than type I errors (42.6% versus 17.0%)

of agents[26]. About 25.6% of the decisions to inspect are unprofitable but have a reward parameter $\vartheta>1$ (Table 3).

(ii) As predicted by the heterogeneous, other-regarding agent model (proposition 3B), spiteful agents are more aggressive inspectors than the one of altruistic agents.

We divided the agents into three groups according to their average investment level in the experiment. We compared the inspecting behavior of high versus low investors while keeping out of the analyses a group of median investors among whom there were no significant differences in individual investment at a 0.05 level. We found that relatively spiteful agents requested on average more inspections per period than relatively altruistic agents did, when controlling for the known level of investment of all the other agents in the group. This conclusion is based on the sign and significance of the coefficient of the dummy variable for highest investors in Table 5 (positive for weak sanctions, negative for strong sanctions) [27].

< *Table 5 about here* >

Although preferences do not account for all the inspection "errors", they play a relevant role. More investigations are needed on the beliefs about other agents' investment levels (see for instance Coats and Neilson, 1999).

9. RESULTS OF STRONG SANCTION EXPERIMENTS

Result 4. In the presence of the strong sanctioning institution agents use the resource more than is predicted by the classical model. Moreover, there is no "reasonable" set of parameters that can be used in the heterogeneous, other regarding agent model that yields levels of resource use as high as those observed in the experiments. In other words, the group outcome is inefficient even though both models predict efficiency above 98% levels.
However, such data suggest that experience might account for the errors in the sense that the models are more accurate when applied to the data from experienced people.

Support: In the two strong sanction experiments, the total group investment was on average 85.1. This level was significantly (0.01 level) higher than the outcome predicted by both the standard model (71.1) and the heterogeneous, other-regarding agent model ([64, 72]). The conclusion does not change when we estimate the Noussair convergence model. The ordinary least squared asymptote is 86.13 (Table 2) and none of the predicted values are in its 95% confidence interval. The group investment across periods ranged from a minimum of 69 to a maximum of 126.
The efficiency level is very high, 93.98%, but still sub-optimal and lower than the target level. When the inspection fees (17.12%) are subtracted, the net rent is 76.87%. Sub-optimality might

---

[26] In our sanction design the reward parameter ϑ increases with the individual investment level of the targeted agent in the interval $x_j \in (\lambda, \underline{x})$. Below λ there is no sanction and above $\underline{x}=\lambda + (k/h)$ inspecting is not costly any more (it is actually profitable and so ϑ is ill-defined). In particular, $\vartheta = \frac{h(x_j-\lambda)}{k-h(x_j-\lambda)}$, if $\lambda < x_j < \underline{x}$. Some numerical values: for weak sanctions, $\vartheta=1$ when $x_j=12.5$, $\vartheta=2$ when $x_j=13.66$ and $\vartheta=4$ when $x_j=14.75$; for strong sanctions, $\vartheta=1$ when $x_j=7.87$, $\vartheta=2$ when $x_j=8.16$ and $\vartheta=4$ when $x_j=8.4$.

[27] Regressions for each single experiment confirm this general conclusion with the exception of one of the weak sanction experiment (0225) where the highest investors dummy is not significant at 0.10 level.

be due to the inexperience of subjects, since there is a significant improvement in the group efficiency when subjects are experienced (gross rent 98.24% versus 89.73%)[28].

Result 5. In the strong sanction environment about 99% of the actions are inspected.
(i) The number of inspections exceeds the number that is profitable. In fact about 50% more inspections take place than is ex-post profitable.
(ii) The lowest investors are more aggressive inspectors than the highest investors
Part (ii) is not predicted by the classical model (Proposition 5A) but is consistent with the heterogeneous, other regarding agent model (Proposition 5B).

Support:

(i) About 99.1% of the actions were inspected, a value definitely close to the 100% predicted[29]. The ratio between the undergone inspections over potentially non-negative balance inspections is 1.49, which is high but less severe than in the weak sanction experiments[30]. All potentially profitable inspections have been requested.

(ii) Relatively spiteful agents (low investors) are more aggressive inspectors than relatively altruistic agents (high investors) since the coefficient of the *highest investor* variable is negative correlated and significantly so with the number of inspections (Table 5).

The functioning of the inspection mechanism under the strong sanction treatment reveals similar features than under the weak sanction treatment. There are too many inspections and spiteful agents are more active inspectors than altruistic agents are.

10.CONCLUSIONS

We replicated in the laboratory two basic outcomes that were found by WGO from the use by a group of people of a common-pool resource in a sanction-free environment:

1. **The `tragedy of the commons' is more severe than was expected.** Cooperation levels are lower than those which self-interested free riders will choose. To be more precise, group efficiency is below the classical Nash equilibrium level (at 39.5% efficiency): -3.2% efficiency in WGO and 29.5% in our experiments. Most of the difference between the two studies occurs in the earlier rounds and dies out over time: the convergence values estimated with the Noussair model are statistically indistinguishable (131.97 WGO and 133.78 ours), although our data reject the Nash equilibrium value of $X=128$ at a 0.05 level where WGO data are more noisy[31].

---

[28] The inequality in the average use across agents is substantially lower than in the no sanction environment (Table 2). The average standard deviation of the agent period earnings from investment is of 5.6 francs and 1.6 francs ones fines are subtracted.

[29] Although almost all the agents were inspected every period, not all the agents requested to inspect everybody every period, as predicted by the classical model. On average an agent requested less than 4 inspections per period instead of 7.

[30] The surprise is the inspection balance that is largely positive. In equilibrium with $(X^*,p^*) = (71.1, 1)$, we expect an inspection balance about 1.4% of the maximum rent. Data talk of an average balance of 18.8%, more than ten times higher than what was predicted. The reason of such "success" was not mainly in the exceptional ability in discovering high investors but in the high average value of total group investment. In fact, about 32.6% of the inspecting decisions were incorrect, which is only slightly lower than in weak sanction. The type II error is very high (0.972) and there are no type I errors.

[31] Average group efficiency is computed using WGO's 3 experiments and the first 20 rounds of the 4 no sanction experiment in this paper. The average group appropriation levels are statistically different at 0.01 level. The 0.95 confidence interval of the Noussair asymptotes are[124.38, 139.56] for WGO and [129.33, 138.22] for ours.

2. **Individual actions are heterogeneous.** At the individual level, the one-shot Nash equilibrium prediction is definitely rejected. In WGO's experiments in the 48% of the rounds not a single agent invested 16 tokens. In our experiments the figure is 90% [32].

We suggest a model with heterogeneous, other-regarding agents that is compatible with the above results. In this model there is room for altruistic, self-interested, and spiteful agents who care to a different degree about other people's earnings. An important role is played by spiteful agents who find enjoyment in decreasing the earnings of others. The same model can explain the outcomes of two variations to the baseline experiments that can be found in the literature:

A. *Restriction of the individual action space.* Since some agents choose actions different from the self-interested best response, an upper bound to the individual appropriation level sufficiently close to the classical equilibrium could improve the group efficiency. WGO shifted the upper bound from 50 to 20 tokens where the individual classical equilibrium was at 16 tokens and reported a substantial group efficiency gain (data have been rescaled).

B. *Introduction of self-administered sanctions.* A system of social sanctions could improve group efficiency because some agents are willing to request sanctions even if it is costly to do so (see Section 3) [33].

In the paper, we report the results of common-pool resource experiments with a specific monitoring and sanctioning system. We tested a weak and a strong sanctioning institution. The rules of the design were modeled after a historical case of use of a common pool resource in the Italian Alps, where people could inspect one another and inflict punishments according to some legal rules. Two of the results are:

1. **Agent heterogeneity in inspection behavior.** Spiteful agents are more aggressive inspectors than altruistic agents and are also willing to request unprofitable inspections. The high rate of unprofitable inspections is also due to mistakes originating from wrong beliefs about the individual appropriation levels of the other agents.
2. **Improvement of the group efficiency.** Under the weak sanction treatment there is a spectacular improvement in efficiency (57.2%) that is not predicted by the classical model. Group behavior in the strong sanction environment shows large improvements (94%) but is not at the optimal level, as predicted by both classical and heterogeneous, other-regarding agent models.

In the explanation of the results under all experimental treatments a key role is played by the heterogeneity of behavior of the agents, and in particular by the presence of few spiteful agents. We have explored the implication of a specific model, but it is likely that other models – maybe as parsimonious as the one suggested here - will be even better supported by the experimental data. We believe, however, that two elements will be pivotal in any explanation: heterogeneity of the agents and some form of other-regarding attitude in the preferences, which could also take the form of reciprocity or fairness.

The self-administered monitoring and sanctioning system studied in this paper performed very well, although it was not perfect, in solving the "tragedy of the commons". In our view, its main strength is the ability of such institutions to turn individual mistakes and socially harmful attitudes into beneficial actions for the group through the possibility of inspections that modify agent incentives. Maintaining an inspection mechanism involves a deadweight loss for the group but a careful choice

---

[32] Part of the increase observed in our experiments might be due to the re-scaling of the action space and to the opportunity to invest any real and not only integer number.

[33] According to the heterogeneous, other-regarding agent model, an improvement in the gross group efficiency will occur when a sanction system is in place in Fehr and Gatcher (1999) (as well as in this paper) because the 'reward' from sanctioning ϑ is lower when agents cooperate than when they free-ride. In Ostrom et al (1994) and in Moir(1998) the conclusion comes only if the agents share a norm of fairness that induce to target especially free-riders.

of the parameters could bring benefits that compensate many times for that loss. In order to understand the consequences on the appropriation system, we suggested using a heterogeneous, other-regarding agent model. Still, there is a need to investigate further the formation of beliefs of agents about others because it plays a crucial role both in appropriation and inspection decisions.

REFERENCES

Beckenkamp, Martin and Axel Ostmann (1999), Missing the target? Sanctioning as an ambiguous structural solution, in: Foddy, Margaret et al (Eds.), *Resolving social dilemmas: dynamic, structural, and intergroup aspects,* Ann Arbor, Psychology Press

Bolton, Gary E. and Axel Ockenfels (1999) ERC: A theory of equity, reciprocity, and competition, *American Economic Review*, forthcoming

Cardenas, Juan Camilo, John Stranlund, and Cleve Willis (1999) Local Environmental Control and Institutional Crowding-Out, *manuscript,* November

Casari, Marco (1997) Governing the commons: a historical analysis of rural villages in the Alps, *Research report, California Institute of Technology, Pasadena,* September

Chan, K.S., Rob Godby, S. Mestelman and R. Andrew Muller (1997) Equity theory and the voluntary provision of public goods, *Journal of Economic Behavior and Organization, 32, 349-64*

Clark, Colin (1976) Mathematical bioeconomics: the optimal management of renewable resources, New York, Wiley

Coats, Jennifer C. and William S. Neilson (1999) Beliefs about altruism, fairness, reciprocity, selfishness and spite, *manuscript,* June

Edgeworth, Francis Ysidro (1881) Mathematical Psychics; an essay on the application of mathematics to the moral sciences, Reprinted in 1967, New York, A.M.Kelley

Fehr, Ernst and Schmidt (1998) A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, forthcoming

Fehr, Ernst and Simon Gächter (1999) Cooperation and Punishment in Public Goods Experiments, *Working paper*, Institute for Empirical Research in Economics, University of Zurich

Gordon, H. 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, 1243-48

Hirshleifer, David and Eric Rasmusen (1989), Cooperation in a repeated prisoner dilemma with ostracism, *Journal of Economic Behavior and Organization*, 12, 87-106

Ito, Masaru, Tatsuyoshi Saijo, and Masashi Une (1995), The tragedy of the commons revisited. Identifying behavioral principles, *Journal of Economic Behavior and Organization*, 28, 311-335

Krebs, Dennis L. (1970), Altruism – an examination of the concept and a review of the literature, *Psychological Bulletin*, 73, 4, 258-302

Ledyard, John (1995) Public goods: A survey of experimental research, in *: The Handbook of experimental economics*, J.H.Kagel and A.E.Roth (Eds.), Princeton University Press, Princeton, N.J., 111-194

Levine, David K. (1997) Modeling altruism and spitefulness in experiments, unpublished, *Department of Economics, UCLA,* October

McCusker, Christopher and Peter J. Carnevale (1995) Framing in Resource Dilemmas: Loss Aversion and the Moderating Effects of Sanctions , *Organizational Behavior and Human Decision Processes*, 61, 2, 190-201

Moir, Robert (1998), Spies and Swords: Costly Monitoring and sanctioning in a common -pool resource environment, version 3, *manuscript*, Dept. of Economics, University of New Brunswick, Canada

Mui, Vai-Lam (1995) The economics of envy, *Journal of Economic Behavior and Organization,* 26, p.311-336

Noussair, Charles N., Charles R. Plott, and Raymond G. Riezman (1995) An Experimental Investigation of the Pattern of International Trade, *American Economic Review,* June, 85, 3, p.462-491

Ostrom, Elinor (1990) Governing the Commons: The Evolution of Institutions for Collective action, New York, Cambridge University Press

Ostrom, Elinor, James Walker, and R. Gardner (1992) Covenants with and without a sword: self-governance is possible, *American Political Science Review*, 86, 404-417

Ostrom, Elinor, Roy Gardner, and James Walker (1994) Rules, Games, and Common-Pool Resources, Ann Arbor, University of Michigan

Palfrey, Thomas R., and Howard Rosenthal (1994) Repeated Play, Cooperation and Coordination: An Experimental Study, *Review of Economic Studies,* 61, p.545-565

Palfrey, Thomas R., and Jeffrey E. Prisbey (1997) Anomalous Behavior in Public Goods Experiments: How Much and Why?, *American Economic Review, Dec*

Rabin, Matthew (1993) Incorporating Fairness into Game Theory and Economics , *American Economic Review*, 83, 5, 1281-1302

Rabin, Matthew (1996) Psychology and Economics, *Working paper,* Department of Economics, University of California at Berkeley, September 28

Smith, Adam (1759) The Theory of Moral Sentiments, Reprinted, Indianapolis, Liberty Classics

Van Winden, Frans, Frans van Dijk, and Joep Sonnemans (1998), Intrinsic motivation in a public good environment, in*: Handbook of experimental economics results*, Charles R. Plott and Vernon L. Smith (Eds.) (forthcoming) North Holland/Elsevier Press

Varian, Hal R. (1990) Monitoring agents with other agents, *Journal of Institutional and Theoretical Economics,* 146, 153-174

Walker, James M., Roy Gardner, and Elinor Ostrom (1990) Rent Dissipation in a Limited-Access Common-Pool Resource: Experimental Evidence, *Journal of Environmental Economics and Management,* 19, 203-211

Walker, James, R. Gardner, Elinor Ostrom, and A.Herr (1996), Voting on allocation rules in a commons: theoretical issues and experimental results, *Workshop in Political Theory and Policy Analysis, W95-18*, Indiana University, Bloomington, Indiana

Yamagishi, T. (1988) Seriousness of social dilemmas and the provision of a sanctioning system , *Social Psychology Quarterly*, 51, 32-42

APPENDIX: PARAMETERS OF THE NO SANCTION DESIGN

Our design:

Number of people in the group: 8 Conversion rate $ 0.04 per 1 franc

Total group earnings (in francs): $\pi_i = 2.5\cdot(E - x_i) + \frac{x_i}{X}\cdot f(X)$

where $x_i \in [0, 50]$ for $i=1, .., 8$; $2.5\cdot(E - x_i)$ is the cost of tokens (endowment E=4 or E=0), and

Gross group return
$$f(X) = \begin{cases} \frac{23}{2}X - \frac{1}{16}X^2, if\ X \le 184 \\ 200\cdot[e^{-0.0575(X-184)} - 1], if\ X > 184 \end{cases}$$

Walker, Gardner, and Ostrom (1990) high endowment design (WGO):

Number of people in the group: 8 Conversion rate $ 0.01 per 1 franc

Total group earnings (in francs): $\pi_i = 5\cdot(25 - x_i) + \frac{x_i}{X}\cdot f(X)$

where $x_i \in [0, 25]$ for $i=1, .., 8$; $5\cdot(E - x_i)$ are the market 1 earnings (endowment always E=25), and

Market 2 earnings $f(X) = 23\cdot X - \frac{1}{4}X^2$

\---------------

There are differences between our no sanction design and WGO both in the incentive structure (1 - 4) and in the way the information is conveyed (A -D) (earnings are expressed in dollars per person per period):

1. The **range of the choice variable x_i has been rescaled** from [0,25] to [0,50] because the difference between two possible equilibrium points, Nash equilibrium and open access, was just one unit. Increasing the perceived number of steps from 25 to 50 might lead to a more accurate investment decision. Moreover, in pilot experiments we noticed that people generally input integer numbers even if the software allowed any real number. This fact may be a problem since in WGO the socially optimal investment level corresponds to a non-integer individual investment (4.5 tokens).
2. The **conversion rate franc/dollar has been increased four times** from $ 0.01 to $ 0.04 per 1 franc in order to maintain a higher effort level by the participants in the experiment. As a result, the *difference* in terms of *individual* earnings between the social optimum and open access points has increased from $ 0.405 to $ 1.62. The adjustment has been suggested by the playful behavior of some subjects during the pilot experiment when the low conversion rate was used.
3. The **minimum safe earning level has been decreased**. If nothing is invested in the "risky" market 2, the original earnings were $1.25. In our setting a zero investment (x_i=0) yields a period return of $0.4 (when E=4). The change implies a downward shift in the payoff but does not affect the incentive structure. The reason of the change is to limit the maximum earnings that would have otherwise been too high given the new conversion rate (point 2).
4. The **Gross group return has been modified** in the interval [184, 400]. The function *f(X)* now has a lower bound at –200 francs that is much higher than it originally was. A group investment of X=184 has an efficiency of –142% in both settings. At X=400 the difference in dollar earnings is small ($ –5.6 instead of $ –6.75). The reason for the change is to limit the maximum loss given the new conversion rate in case people "go crazy" (see point 2). In the experiments conducted, we observed an investment level above 184 just in one round out of 291.

< Table 1A about here >

A. **Graph instead of formula**. WGO provided the subjects with the analytical expression of the gross group return from market 2. The expression may have been used to compute the equilibrium. We replaced it with a plot of the function f(X) because we think that the graph would easy the understanding of the basic underlying phenomenon. The graph in the instructions is similar to the upper part of Figure 1 once the cost line is removed.

B. **Detailed table.** In order to compute the equilibrium, subjects could use a very detailed table of gross group return. The table gives the gross group return and the return per token invested for 100 values of the total group investment, compared to the 10 values given by WGO. All theoretical equilibrium points are listed in the table given to the subjects. Our table does not supply the marginal returns, which instead WGO provided.

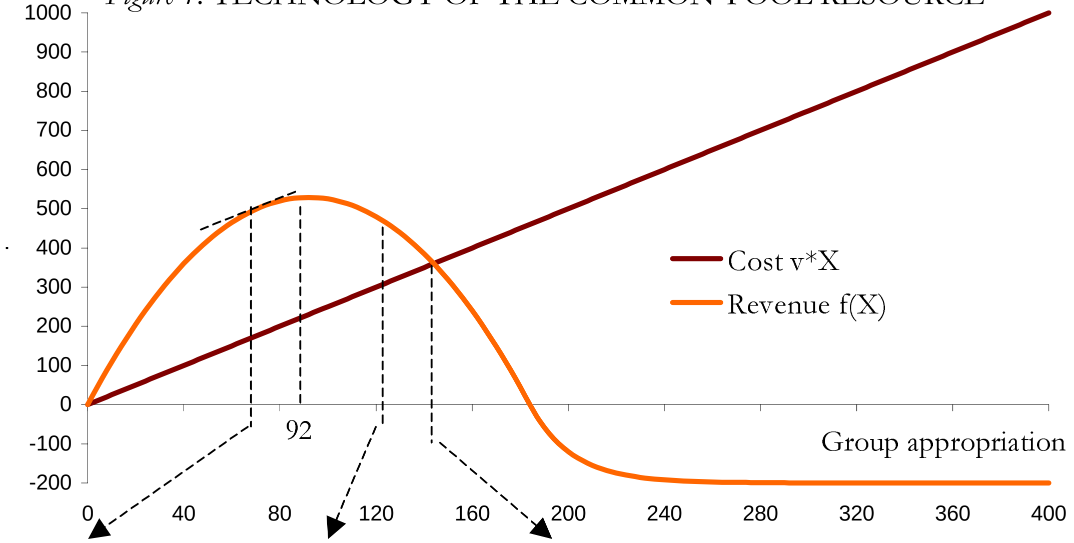
C. **Different software**. The software was run on Netscape and was written specifically for this application. It includes a calculator to compute the cost of tokens, the gross group return , and the individual share of gross for every possible real level of investment in the admissible range.

D. In WGO **market 1** represents the *opportunity cost* of the investment and yields a constant return. The way in which it was presented to the subjects in this study is as a *direct cost* of the investment. You can order the tokens to invest and pay a constant unit price for them. This change may make the decision of the agents easier by suggesting a comparison of the price of tokens with the return from the market.

< Table 2A about here >

FIGURES and TABLES:

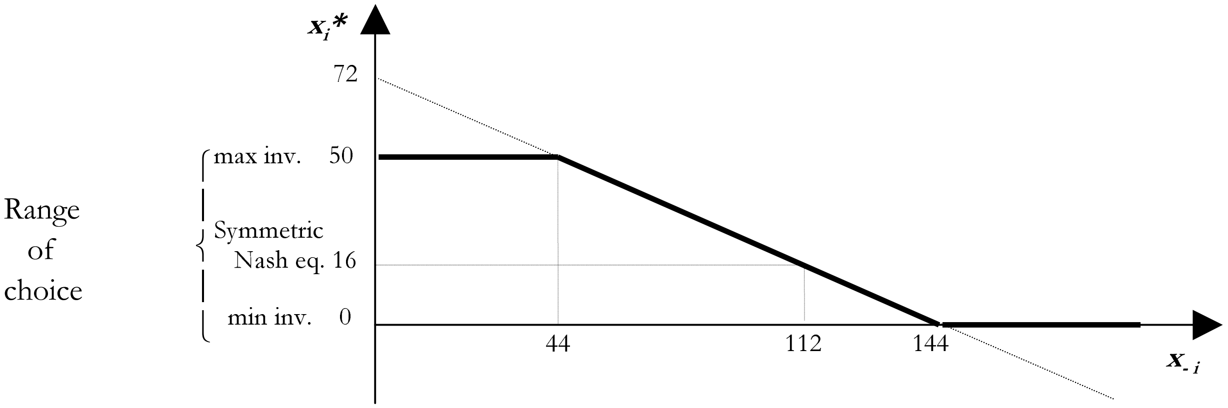
Figure 1: TECHNOLOGY OF THE COMMON-POOL RESOURCE



OPTIMAL USE	*NASH EQUILIBRIUM*	OPEN ACCESS	
72	128	144	Group appropriation X
9	16	18	Symmetric individual use
100%	39.5%	0%	Efficiency level *(% of maximum rent)*

Notes: The group welfare is W= f(X) - v·X, (the endowment E is set to zero); efficiency increases in [0,72) and decreases in (72,400]

Figure 2: BEST RESPONSE FUNCTION, no-sanction treatment



Notes: In the no-sanction design (N=8, a=23/2, v=2.5, b=1/16) the best response function is $x_i^* = 72 - \frac{1}{2} x_{-i}$ *(bold line)*

Figure 3: AVERAGE EFFICIENCY BY EXPERIMENT

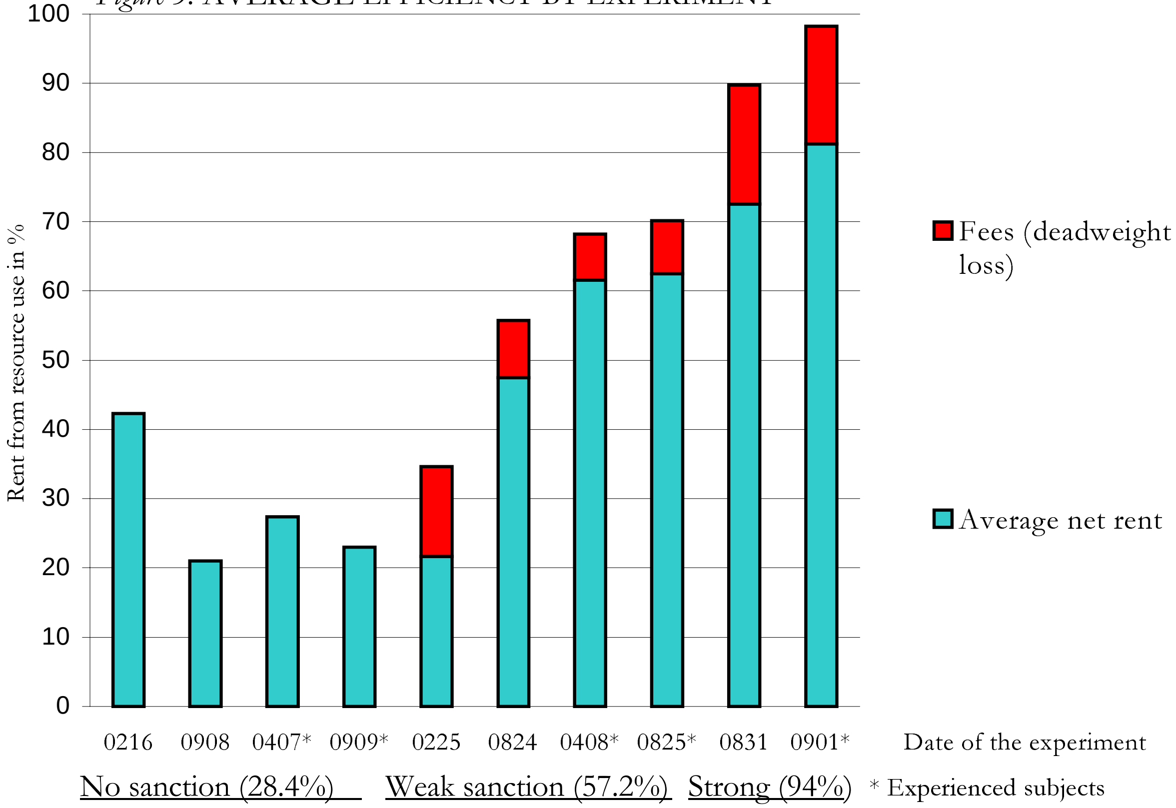


Figure 4: AVERAGE GROUP APPROPRIATION BY EXPERIMENTAL TREATMENT

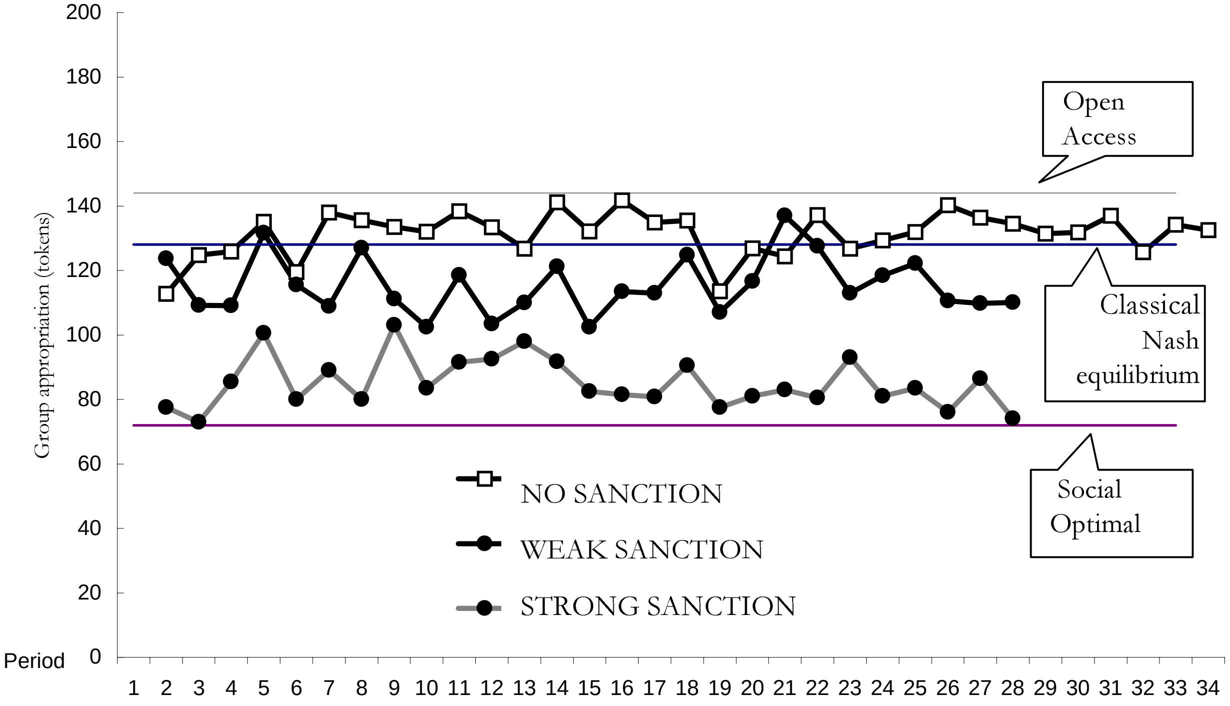
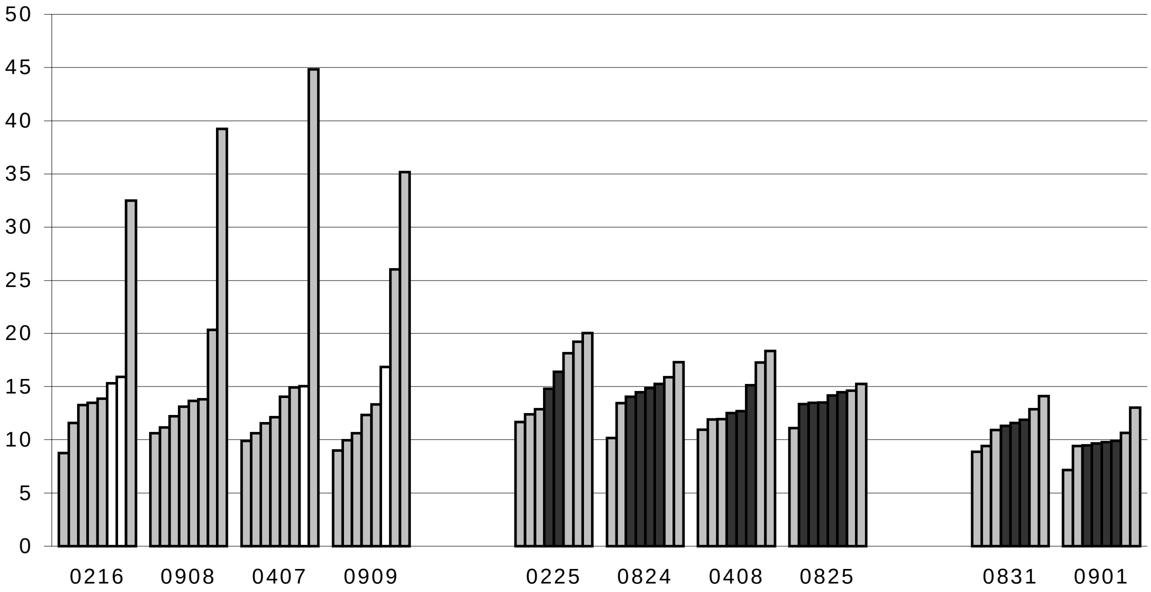


Figure 5: AGENT AVERAGE APPROPRIATION LEVELS (tokens)



Notes: Compare to the individual classical Nash equilibria: x_i*=16 for no sanction and weak sanction designs and* $x_i \approx 8.9$ *for strong sanction design. The white bars in the no-sanction experiments indicate the agents whose behavior is not statistically different from the classical Nash equilibrium at 0.05 level. The darker colored bars in the sanction experiments indicate the group of median agents that are not significantly different one from the other at 0.05 level.*

Table 1: EXPERIMENTAL DESIGN OF MONITORING AND SANCTIONING DEVICES

	CASARI and PLOTT (1999)	**MOIR (1998)**	**OSTROM et al (1994)**	**FEHR and GATCHER (1999)**
Environment:	**Common-pool Resource**	Common-pool Resource	Common-pool Resource	Public Goods Provision
MONITORING				
• *Fee*	**Yes, fixed fee for each request**	Yes, variable fee for each request	No	No
• *Are all individual actions revealed?*	**No, only if somebody in the group requests it**	No, only if the agent requests it	Yes, automatically and classified by ID	Yes, automatically and anonymously
SANCTIONING				
Targeted agent:				
• *Amount of the fine*	**In a fixed proportion of over-use**	At the will of inspector (variable upper bound)	At the will of inspector (fixed upper bound)	At the will of inspector (up to 100% of period earnings)
• *Condition for inflicting the fine*	**If over-use occurred**	If over-use occurred	None	None
• *Multiple fines on the same action*	**No**	Yes	Yes	Yes
• *Identity of targeted agent*	**Public**	Public	?	?
Requesting agent:				
• *Fee*	**No, already paid for monitoring**	Proportional to the amount of the sanction	Proportional to the amount of the sanction	More than proportional to the amount of the sanction
• *Who receives the fine*	**Inspector**	Experimenter	Experimenter	Experimenter
• *Limits to requests of sanctions per period*	**No**	Yes, the individual period profits must cover sanction fees	Yes, each agent can do only a single request	?
• *Identity of requesting agent*	**Not revealed**	Not revealed	Not revealed	Not revealed?

Notes: Monitoring is always perfect (i.e. there is truthful revelation of the action).
In Ostrom(1992) and Fehr and Gatcher (1999) agents can sanction each other but there is really no monitoring device, since the individual actions are automatically revealed to everybody at the end of each period. Moir(1998) introduces two distinct decisions, first to monitor an agent and then to eventually sanction her. We have compacted them in a single decision: to inspect an agent or not. An inspection uncovers another agent's action and automatically inflicts a sanction if some conditions are met.

Table 2: SUMMARY TABLE FOR APPROPRIATION DECISIONS

Experiments	NO SANCTION				WEAK SANCTION				STRONG SANCTION	
Date	0216	0908	0407	0909	0225	0824	0408	0825	0831	0901
Sanctions	No	No	No	No	Yes	Yes	Yes	Yes	Yes*	Yes*
Experience	No	No	Yes	Yes+	No	No	Yes	Yes+	No	Yes+
Number of rounds	32^	32	33	32	27	27	27	27	27	27
Period endowment (tokens)	4	4	4	0	4	4	4	0	0	0
Conversion rate ($ per franc)	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.03
GROUP APPROPRIATION										
Average	124.71	134.17	133.08	133.31	125.60	115.44	110.81	110.04	91.07	79.07
Classical Nash equilibrium	*128*	*128*	*128*	*128*	*128*	*128*	*128*	*128*	*71.11*	*71.11*
Minimum	100	85.5	121.5	87	87.3	87	92.5	90	74	69
Maximum	154	161	149.5	167	186.8	160	154.7	137	126	95.5
Standard deviation (Sd)	14.84	15.49	6.01	15.46	23.13	20.57	12.15	10.20	13.22	6.50
First half Sd/Second half Sd	4.52	1.68	2.59	0.79	0.92	0.83	2.57	1.46	2.18	0.85
EVOLUTION OF GROUP APPROPRIATION OVER TIME (Noussair model)										
- Starting point	105.69	97.72	122.37	131.90	147.86	106.02	118.44	105.88	93.79	63.85
- Asymptote	133.96				114.78				86.13	
[0.95 confidence interval]	[130.96, 136.57]				[110.51, 119.06]				[82.23,90.02]	
GROUP EFFICIENCY (% of maximum rent)										
Average Rent	42.29	20.97	27.36	23.00	34.67	55.74	68.20	70.15	89.73	98.24
First 25 periods	*42.48*	*23.74*	*27.13*	*22.27*	*32.43*	*52.71*	*68.94*	*70.75*	*89.29*	*98.14*
Last 2 periods (after announce.)	*54.23*	*1.28*	*28.21*	*41.06*	*62.68*	*93.58*	*58.95*	*62.62*	*95.29*	*99.47*
INDIVIDUAL AVERAGE APPROPRIATION LEVELS										
Lowest average investor	8.8	10.6	9.9	9.0	11.7	10.2	11.0	11.1	8.9	7.1
Highest average investor	32.5	39.2	44.8	35.2	20.0	17.3	18.4	15.3	14.1	13.0
Gini coefficient of average agent appropriations	0.195	0.262	0.234	0.265	0.109	0.099	0.072	0.044	0.078	0.077
Rank correlation 1st/2nd half	0.917	0.902	0.961	0.966	0.907	0.892	0.926	0.872	0.823	0.858

Notes:

Date	Experiments were done at the California Institute of Technology in 1998
Sanctions	"No" is a no sanction experiment; "Yes" means that a monitoring and sanctioning device was added to the no sanction experiment; "Yes*" indicates a different set of sanctioning parameters which affected the equilibrium point (see paragraph 5.3 for details)
Experience	"No" means that no subject has ever participated in this type of experiment before; "Yes" means that all the subjects have already participated in this type of experiment (on Dec 9, 1997 at the earliest); "Yes+" means that all the subjects have participated the day before in this type of experiment with the same group of people
No. of rounds	Number of effective rounds of play, which excludes two practice rounds; (^) On 0216 a paper copy of the Return from investment table was handed to the subjects between the 10th and the 11th round instead of before the 1st round.

Period endowment differs among experiments for the only purpose of keeping the cost of the experiments under control. Sd=Standard deviations; Noussair model: three ordinary least squared regressions, one for each experimental treatment.

Table 3: APPROPRIATION AND INSPECTION ACTIONS

		WEAK SANCTION				STRONG SANCTION			
		Inspected?				Inspected?			
		No	Yes	*Totals*		No	Yes	*Totals*	
CAN YOU INSPECT THE ACTION WITH PROFIT?	No, $\vartheta<1$	220	162	*382*	*44.2%*	2	108	*110*	*25.5%*
	No, $\vartheta\geq1$	152	114	*266*	*30.8%*	2	33	*35*	*8.1%*
	Zero	20	37	*57*	*6.6%*	0	0	*0*	*0.0%*
	Yes	27	132	*159*	*18.4%*	0	287	*287*	*66.4%*
	Totals	*419*	*445*	*864*	*100%*	*4*	*428*	*432*	*100%*
		48.5%	*51.5%*	*100%*		*0.9%*	*99.1%*	*100%*	

Notes: No agent will request an inspection with a reward parameter $\vartheta<1$ in both the classical and other-regarding agent models. The reward parameter ϑ is below 1 when $x_j<12.5$ with weak sanctions and $x_j<7.875$ with strong sanctions. Inspecting is profitable when $x_j>16$ with weak sanctions and when $x_j>9$ with strong sanctions.

Table 4: SUMMARY TABLE FOR INSPECTION DECISIONS

Sanction Experiments	*WEAK*				**STRONG**	
Date	0225	0824	0408	0825	0831	0901
WELFARE ANALYSES	**(% of the maximum rent)**					
(1) Fees (deadweight loss)	13.04	8.24	6.64	7.68	17.20	17.04
(2) Fines (transfers)	14.01	8.90	6.49	6.19	43.39	28.44
Net Rent (Rent – Fees (1))	21.63	47.49	61.56	62.47	72.53	81.20
Inspection balance ((2) – (1))	1.0	0.7	-0.1	-1.5	26.2	11.4
ANALYSIS BY ACTIONS						
Actions inspected on total	75.5%	47.7%	38.4%	44.4%	99.5%	98.6%
Ratio of inspections undergone over the number of zero or positive balance inspections	2.17	2.29	1.48	2.40	1.52	1.46
Inspection “errors” on total actions	47.7%	34.3%	25.9%	32.4%	34.3%	31.0%
Of which:						
Type I error *(share of potentially profitable inspections not requested)*	0.098	0.167	0.239	0.192	0.000	0.000
Type II error *(share of potentially non-profitable inspections that have been requested)*	0.695	0.398	0.281	0.369	0.987	0.957

Notes: Total number of actions in an experiment: 216; maximum rent from investment for the group is 324 francs per period; Balance of an inspection is defined from the standpoint of the agent who asked to inspect: fine collected minus fee paid. Inspection "errors" is in quote because strictly speaking they are mistakes only from the point of view of a self-interested agent.

Table 5: SPITEFUL AGENTS INSPECT MORE THAN ALTRUISTIC AGENTS

	WEAK sanctions		**STRONG sanctions**	
OLS regression	Coefficient	p-value	Coefficient	p-value
Dependent variable:	Total number of requests of inspections per period			
Sample size without median investors:	486		243	
Independent variables:				
Highest investors (dummy)	0.34	0.015	-1.23	0.000
Period investment of the other agents	0.06	0.000	0.03	0.159
Constant	-5.42	0.000	1.58	0.366

Notes: the classical model predicts insignificant coefficients for the highest investors dummy variable. See Figure 5 to identify the median investors whose actions were excluded from the regressions.

Table 1A: COMPARISON BETWEEN OUR DESIGN AND THOSE OF WGO

	OUR DESIGN (E=4)				**WGO** (1990)			
	SO	NASH	OA	MAX	SO	NASH	OA	MAX
Total group investment (tokens)	72	128	144	400	36	64	72	200
Symmetric individual investment (tokens)	9	16	18	50	4.5	8	9	25
Individual earnings per period (in francs)	50.5	26	10	-140	165.5	141	125	-675
Cost of tokens (in francs)	-12.5	-30	-35	-115	102.5	85	80	0
Gross Return (in francs)	63	56	45	-25	63	56	45	-675
Individual earnings per period ($)	2.02	1.04	0.4	-5.6	1.655	1.41	1.25	-6.75

Notes: SO=social optimum, NASH=classic Nash equilibrium, OA=open access, MAX= maximum investment

Table 2A: SUMMARY TABLE FOR WGO

Experiments	**NO SANCTION**		
Experiment number	1	2	3
Experience	Yes	Yes	Yes
Number of periods	20	20	20
Period endowment*	50	50	50
Conversion rate ($ per franc)	0.01	0.01	0.01
GROUP INVESTMENT			
Average	138.3	147.5	136.6
Classical Nash equilibrium	*128*	*128*	*128*
Minimum	108	84	116
Maximum	188	230	176
Standard deviation (Sd)	23.25	31.59	18.30
First half Sd/Second half Sd	1.58	2.22	1.56
GROUP EFFICIENCY	*(% of*	*maximum rent)*	
Average Rent	5.30%	-28.24%	13.36%

Notes: *Original data have been rescaled (i.e. multiplied by two) to make the comparison easier with Table 3.*