THE PERILS OF PEER PUNISHMENT. EVIDENCE FROM A COMMON POOL RESOURCE FRAMED FIELD EXPERIMENT

Gioia de Melo y Matías Piaggio

Octubre 2012

INSTITUTO DE ECONOMIA
Serie Documentos de Trabajo
DT 16/12

UNIVERSIDAD DE LA REPÚBLICA (UDELAR)- FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN- INSTITUTO DE ECONOMÍA (FCEYA)

URUGUAY

ISSN: 1510-9305 IMPRESO

ISSN: 1688-5090 ONLINE

THE PERILS OF PEER PUNISHMENT. EVIDENCE FROM A COMMON POOL RESOURCE FRAMED FIELD EXPERIMENT

Gioia de Melo[1] and Matías Piaggio[2]

We provide a model and experimental evidence on the effects of non-monetary punishment (NMP) by peers among communities of Uruguayan fishers exploiting a common pool resource (CPR). We find a) experimental groups composed of fishers from different communities (out-groups) who are sometimes in conflict over fishing territories did not overexploit the resource more than groups from a single community (in-groups) and, unlike in-groups, out-groups reduced their exploitation of the resource in response to the threat of punishment; b) cooperative individuals punished free riders while a substantial amount of punishment was targeted by free riders on cooperators, who [in turn] responded by increasing their exploitation of the resource; and c) wealthier individuals practiced greater overexploitation of the resource. Our results suggest that the relevance of in-group favoritism in promoting cooperation due to social preferences may be overrated, and that the effectiveness of peer punishment is greater when individuals are motivated by social preferences and also that coordination is required to prevent anti-social targeting and to enhance the social signal conveyed by the punishment.

JEL: D03, O12, C93

Keywords: non-monetary punishment, in-group bias, frame field experiment, social preferences, common pool resource

[1] Instituto de Economía, Facultad de Ciencias Económicas y de Administración, Universidad de la República y Dipartimento di Economia Politica -Università Degli Studi di Siena.

[2] Instituto de Economía, Facultad de Ciencias Económicas y de Administración, Universidad de la República y Departament d'Economia Aplicada - Universitat Autònoma de Barcelona.

1 Introduction

The conservation and sustainable use of common pool resources (CPRs) is an important issue worldwide not only because of the relevance that conservation of biodiversity has, but also because local commons are essential for the livelihoods of the world's poorest communities. The exploitation of a CPR poses a typical social dilemma. Hardin (1968) proposed to establish either private or state property rights as a solution to avoid the tragedy of the commons. However, market contracts and governments often fail to prevent overexploitation because the necessary information to design and enforce beneficial exchanges and directives cannot be effectively used by judges and government officials (Bowles and Gintis, 2002). In the last decades, many authors have argued that communal property regimes may enhance cooperation in the preservation of a CPR by enforcing social norms, and in this way fill the gaps of incomplete contracts (Ostrom, 1990; Feeny et al. 1990; Baland and Platteau, 1996; Ostrom et al., 1999 and Ostrom, 2000; Bowles and Gintis, 2002).

Several studies have concentrated on the determinants of successful experiences based on communal property regimes but the issue is far from settled.[3] It has been widely argued that a self-sustaining cooperative equilibrium can be achieved in a context in which self-regarding agents interact in repeated games. But in addition to reasons for cooperating related to the expectation of a subsequent reciprocal monetary benefit sufficient to offset the current cost (Trivers, 1971), social preferences may also constitute a motive for cooperation. Social preferences encompass a wide range of motives such as reciprocity, altruism, conformism and also emotions such as shame, guilt and anger (Bowles and Gintis, 2011). In a CPR context, Velez et al. (2009) show that what induces cooperation among Colombian artisanal fishing communities is their desire to conform to or emulate others' behavior in the group. In a similar framework, López et al. (2012) argue that shame can promote cooperative behavior.

In their seminal paper, Fehr and Gätcher (2000) show that individuals are willing to bear a cost in order to punish free riders in public good games. This pattern has also been observed in the context of a CPR dilemma (van Soest and Vyrastekova, 2006; Noussair et al., 2011). However, while van Soest and Vyrastekova (2006) find that costly punishment is effective in increasing cooperation, Noussair et al. (2011) do not observe significant changes in cooperation. Janssen et al. (2010) conclude that costly punishment is not effective in reducing extraction unless combined with communication. There is also evidence that non-monetary punishment (NMP) (Masclet, 2003; Noussair and Tucker, 2005, Dugar, 2010), social approval (Gächter and Fehr, 1999) and public observability (Barr, 2001; Denant-Boemont, 2011; López, 2012) can be effective for increasing contributions in public good games. Rege and Telle (2004) and Noussair and Tucker (2007) show that initial increases in cooperation as a consequence of public observability tend to

[3] For a description of successful cases see Ostrom (1990), Feeney et al. (1990), Ostrom et al. (1999) and Baland et al. (2007).

fade away in a repeated game context. Due to the absence of monetary incentives, non-monetary punishment allows better isolation of the presence of pro-social emotions when an individual reacts to being punished relative to costly punishment. Assessing the relevance of non-monetary punishment as a tool to enhance cooperation is of particular importance in regard to community management of common pool resources, because informal sanctions typically take place in that setting. Nevertheless, to our knowledge, in the context of a CPR dilemma there is so far no evidence of the effectiveness of non-monetary punishment in promoting cooperation.

Several studies that range from laboratory to field experiments show that individuals may achieve greater levels of cooperation when interacting with members of their own group rather than with outsiders (Bandiera et al., 2005; Miguel and Gugerty, 2005; Ruffle and Sosis, 2006; Goette et al. 2006; Bernhard, et al., 2006; Chen and Xin, 2009).[4] Miguel and Gugerty (2005) argue that in rural western Kenya, social sanctions levied against free riders are more likely to take place within ethnic groups, due to the mutual reciprocity that exists within groups, rather than across groups. Bernhard, et al. (2006) observe in a third-party punishment game with two small native groups from Papua New Guinea that the punishment is harsher if the victim of a norm violation belongs to the same group as the punisher. These authors show that in situations in which dictators expect the same level of punishment, they give larger transfers to in-group members than to out-group members. This would point to the relevance of unconditional preferences such as altruism that vary depending on group membership.

In this study, we assess whether non-monetary punishment (NMP) is effective in promoting cooperation via social preferences in a CPR dilemma.[5] We are also interested in determining whether social preferences, such as altruism, reciprocity, shame and anger differ in a context in which individuals exploiting a common pool resource belong to different communities relative to the case in which only individuals from the same community are allowed to exploit the resource. In accord with Bernhard et al. (2006), Goette et al. (2006), and Tanaka et al. (2008), we test in-group favoritism in naturally-occurring groups in field settings. That is, we test whether fishermen that live in different communities have a greater propensity to cooperate when they interact among themselves, than when interacting with fishers who do not belong to their community, and we explore whether they differ in their sensitivity to NMP in these two scenarios.

[4] In a broader sense, Akerlof and Kranton (2000; 2005) and Bowles and Gintis (2002) have highlighted the relevance that social identity and group affiliation have on individuals' behavior in most economic organizations.

[5] For the purpose of this study we define cooperation in a narrow sense as the behavior through which one agent internalizes some of the externalities he imposes on other users, and maintains his own use below what would maximize his individual profits (Baland et al., 2007). Baland et al. (2007) note that cooperation often requires coordination. That is, the creation of institutions is needed in order to regulate the use of the resource. In this study we concentrate on the simplest form of cooperation, as the experiment does not allow for communication or the introduction of any institutional form.

We perform a framed field experiment (Harrison and List, 2004) where the subject pool is fishers from the Uruguayan sea coast who fish in two coastal lagoons and live in nearby villages. We concentrate on coastal lagoons because, unlike the open sea where large-scale fishing is widespread, in coastal lagoons the only agents who develop fishing activities are artisanal fishermen. Fishermen from different communities do not interact during their daily life, but they are used to encountering each other while fishing, as they tend to move from one lagoon to the other depending on fish availability. We implement both an in-group/out-group treatment and a NMP treatment. During the in-group treatment subjects played a CPR game only with members of their own community, while during the out-group treatment we required that they play the game with members of another community. The NMP implied that by facing a monetary cost, individuals could express their disapproval of others' extraction decisions. Disapproval was reflected by receiving flags that vary in color in accordance with the level of disapproval achieved among the rest of the group members.

To our knowledge, this is the first study that examines the relevance of informal sanctions in a CPR game. Additionally, our study combines three innovative features that have not been implemented at the same time before. First, instead of inducing artificial in-group/out-group differences we enable individuals from different communities meeting each other. Second, groups are reshuffled after each period in order to avoid repeated game effects that could lead to a self-sustaining cooperative equilibrium. Third, individuals are charged a monetary cost for punishing others even if those socially punished do not face any monetary cost. This step was implemented in order to avoid subjects punishing the others carelessly.

We find no evidence of in-group favoritism. On the contrary, interacting with fishers from other communities has a positive effect on cooperation when punishment is available. That is, during the out-group treatment, individuals reduce their extraction level when NMP is available irrespective of whether they are effectively punished. We observe that the effectiveness of informal sanctions deteriorates by the fact that not all individuals' types are sensitive to NMP, and that these types of sanctions can be used to punish both free riders and cooperators. We argue that for peer punishment to be effective it should require coordination to prevent anti-social targeting and to enhance the social signal conveyed by the punishment. Finally, we observe that individuals adjust their extraction levels period by period according to their deviation, with respect to the group's average in a previous period, as if following an implicit social norm.

The paper is organized as follows: Section 2 develops a theoretical model; Section 3 describes the experimental design; Section 4 reports results; finally, Section 5 concludes.

2 Theoretical framework

In this section we suggest a model that disentangles different motives of behavior that individuals may face when interacting in the context of a common pool resource dilemma. The model is an adaptation of a model developed by Bowles (2004), in which peer monitoring and forms of social disapproval enable individuals to achieve agreed levels of effort.

In this model individuals not only care about their own payoffs but also value (either positively or negatively) the material payoffs of their peers. Individuals may experience spite or altruism (which are independent of the others' actions), as well as reciprocity (the value they place on others' payoffs depends on the others´ past behavior or in their beliefs about the others' types). Individuals share a social norm regarding how much extraction is admitted and may experience shame if they are publically sanctioned for violating it. Besides, they face motives for punishing others socially when the others deviate from the social norm. However, punishing others may be costly, as it can deteriorate the relationship with ones' peers.

Consider a common pool resource exploited by two individuals *i* and *j* (the model can easily be generalized to *n* members). We define individual *i*'s utility function as:

$$(1) \qquad u_i = \Pi_i + \beta_{ij}\Pi_j - s_i + d_i$$

Utility is the sum of the individual's own material payoffs (Π_i) plus the valuation of the others' material payoffs ($\beta_{ij}\Pi_j$), minus the subjective valuation of shame (s_i), and subjective utility or relief experienced from expressing NMP (d_i).

Let *i*´s material payoff be:

$$(2) \qquad \Pi_i = p_1 a_i + p_2 \left[\sum_{k=i}^{k=j} (a^{Max} - a_k)\right] \ \forall\, k = i, j \text{ and } i \neq j$$

where both *i* and *j* have the same maximum endowment of a^{Max} to use for extracting a particular resource (i.e. fishnets). As in Cardenas (2004), the aggregate extraction by the two players $\left[\sum_{k=i}^{k=j} a_k\right] \forall\, k = i, j \text{ and } i \neq j$ reduces *i*'s material payoffs. Alternatively, the externality can also be described as a public good benefit from conservation, i.e. lack of extraction.

Individuals have preferences as to the other's payoffs. This is reflected in *i*'s utility by the coefficient β, which depends on both unconditional preferences (altruism or spite) and on reciprocity. Member *i*'s degree of other regarding preferences towards *j* is:

$$(3) \qquad \beta_{ij} = \alpha_i + \lambda_i\left(b - a_j\right)$$

where $\alpha_i \in (-1,1)$ is *i*'s unconditional spite or altruism, and $\lambda_i \in (0,1)$ is her degree of reciprocity. The level of reciprocal motivation therefore depends on the extent to which *j* has deviated from the social extraction norm (*b*): If *j* has extracted less than b, and $\lambda_i > 0$ then *i* experiences good will toward *j* and positively values his payoffs. But if *j* extracted more than the social extraction norm or if *i* feels spite towards *j*, then *i* may experience malevolence toward *j* ($\beta_{ij} < 0$) and enhance his utility by disapproving of *j*'s performance.

Punishment for deviating from the social norm is expressed through non-monetary mechanisms. Socially punishing defectors enhances one's utility through the relief of expressing emotions.[6] But it does not come without cost. Individuals who express their disapproval in relation to others' actions face a cost of $c(\mu) = \frac{\gamma_i}{2}\mu_{ij}^2$ due to a deterioration in the relationship with their peers. Therefore, disapproval motives (d_i) are a function of the benefits that punishing provides ($-\beta_{ij}\mu_{ij}$), which is a function of the punishment provided as well as altruism (spite), reciprocity, how the other deviates from the social norm, and the cost of punishing:

$$(4) \qquad d_i = \mu_{ij}\left[-\beta_{ij} - \frac{\gamma_i}{2}\mu_{ij}\right]$$

Finally, being publicly punished by others may cause shame,

$$(5) \qquad s_i = \sigma_i(a_i - b)\mu_{ji}$$

where σ_i is a measure of one´s susceptibility to social punishment. The level of disutility experienced from being socially punished depends on one's susceptibility to punishment, the degree of divergence of the individual's extraction relative to the social norm (how fair the individual evaluates the punishment as being) and the severity of the punishment received.

[6] We are ruling out dynamic effects of punishment, that is, punishing *j* in order to get him to cooperate in the future.

The individual chooses an extraction level a_i and a level of NMP μ_{ij} toward his peer in order to maximize equation (1). The first order condition indicates that the utility-maximizing level of punishment is:

$$(6)\ \mu_{ij} = -\frac{\beta_{ij}}{\gamma_i} = -\frac{1}{\gamma_i}\left[\alpha_i + \lambda_i(b - a_j)\right]$$

In other words, i chooses the level of disapproval that equates the marginal cost of punishment ($\gamma_i\mu_{ij}$) with the marginal benefit of punishment, $-\beta_{ij}$, the negative of the valuation placed on the payoff of the other (as long as $\beta_{ij} < 0$, and chooses zero punishment otherwise). Where punishment is positive, it is clearly increasing in λ_i and decreasing in α_i, as one would expect.

The extraction level of i will be a function of j´s level of extraction, as well as of the parameters related to the other regarding preferences.

$$(7)\ a_i = \frac{\gamma_j\left(p_1 - p_2\left(1 + \beta_{ij}\right)\right)}{2\lambda_j\sigma_i} + b + \frac{\alpha_j}{2\lambda_j}$$

Eq. 7 suggests that i's extraction level varies positively the more altruist j is, the higher j's marginal cost of disapproving is and the higher j's extraction levels are. In turn, i's level of extraction will diminish if he is very sensitive to shame, or if j is a reciprocator. In this way, social preferences (other than spite) may induce individuals to behave in ways that diverge from the Nash equilibrium in a social dilemma. That is, by valuing other players' payoffs, altruism and reciprocity can make individuals behave closer to the social optimum. Reciprocity motives may also induce an individual to express NMP to norm violators. If individuals feel shame when punished by others, this may also help avert coordination failures in terms of resource extraction.

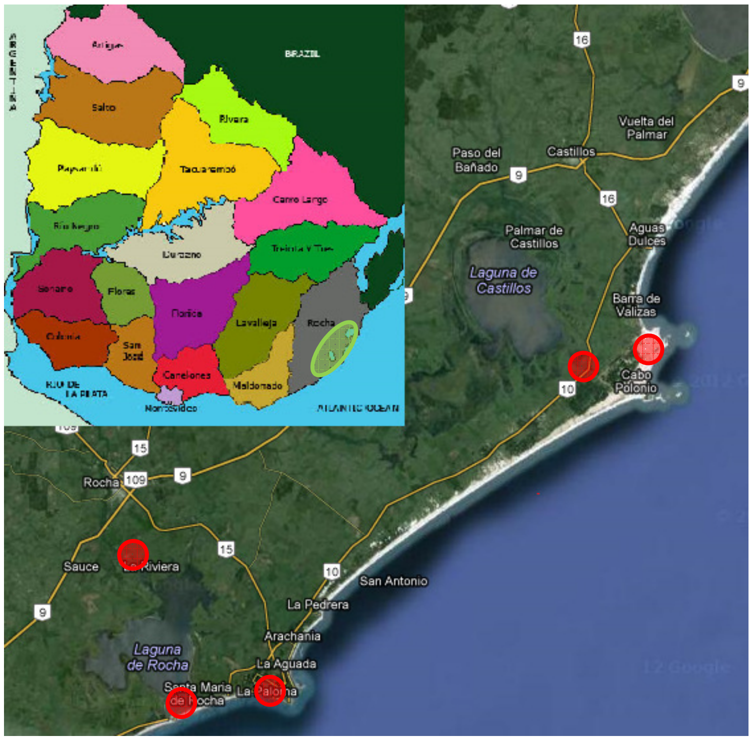
3 Experimental design

3.1 Subject pool

Fishermen from five communities that fish either in the Laguna de Rocha or in Laguna de Castillos (two coastal lagoons 50 kilometers away from each other on the Uruguayan sea coast) or in both were recruited (Figure 1). We consider a community to be a group of people that live in the same settlement and constantly interact among each other. Individuals from different communities do not differ in terms of ethnicity, while they show some differences in socioeconomic characteristics. These communities differ in terms of how connected they are to the rest of society and the exit options they face. While some live very isolated and fishing is their main source of income (Laguna de Rocha, Puerto los

Botes and to less extent El Puente), others are more connected to more densely populated areas and can exploit other exit options (Valizas and Barrio Parque). Facing other exit options as therefore reflected in their income and wealth (see Table A.1 in the Appendix).

Figure 1: Location of field experiment (the five communities marked by red dots)



Fishermen from different communities are not used to meeting each other in their daily lives, but they do so when they move across lagoons during fishing high seasons. This is particularly important during the shrimp high season, which usually takes place once a year in the Laguna de Castillos, but rarely in the Laguna de Rocha, because of geographical reasons. PROBIDES (2002) reports that fishermen complained about fishermen from other communities coming in the high season to fish in the lagoon where they, the complainants, fish all year round. We believe that the place of residence is one of the main dividing factors among fishermen from different communities.

3.2 The experiment

The experiment consisted of a CPR game of 20 periods, structured in two stages of 10 periods each. In each stage, players interacted either only with members of their own community (in-group treatment) or in groups mixed with fishermen from another community (out-group treatment). This was not explained to the subjects. By this we mean that we did not mention during the in-group stage that all members from the same community were going to play together. Individuals were simply told in which group they would play based on their identifier.

The CPR game was framed around the decision of how many nets to use when fishing. Subjects made their decision in subgroups of four subjects. During the first five periods of each stage subjects played a regular CPR game, in which they consider a common pool resource exploited by individuals who have the same maximum endowment (eight nets) to fish. Individual benefits increase in the number of nets one uses and decrease with the aggregate level of nets used (see Table A.2 in the Appendix). Player-i's earnings in each period during the first five periods of each stage were given by the payoff function:

$$\pi_i = 18a_i + 12\sum_{j=1}^{4}(8 - a_j)$$

During the last five periods of each stage a NMP treatment was conducted. In this treatment subjects were allowed to express disapproval of others' fishnet choices. As a consequence, subjects who were punished by other players were assigned a flag, and its color (yellow, orange, or red) indicated how much their peers disapproved of the number of nets they had decided to throw. After making the usual decision on how many nets to use and being informed of the total number of nets used by the subgroup (and therefore being able to determine the others' average extraction), they were able to allocate 0 to 10 disapproval points to each possible choice of fishnets the others may have made (see Table A.3 in Appendix).[7]

Punishment points implied no monetary cost to the punished but did imply a monetary cost to the punisher. The cost of each punishment point for the punisher was equivalent to one point in his earnings account. The subject was charged for the total number of disapproval points he used to punish others, irrespective of whether someone had actually chosen the number of nets the punisher decided to punish.[8] In this way, player i's payoff function during the last five periods of each stage is:

$$\pi_i = 18a_i + 12\sum_{j=1}^{4}(8 - a_j) - \sum_{k=1}^{8}\mu_{ij,k}$$

The cost of punishing was set quite low compared to the points a subject could earn during one period. For instance, in one period if all subjects played the Nash equilibrium, each would earn 144 points, whereas if the social optimum was achieved each would earn 354 points. If during the NMP treatment the subject decided to disapprove of all possible fishnet choices with the maximum number of disapproval points, his cost would amount to 80 points (0.5 US dollars). The aim of this treatment was to recreate the state of being socially

[7] Subjects could also choose to punish a choice of a number of nets identical to their own. In that case, the punishment would be directed solely to others and not to themselves. This was only explained in case someone asked. Potentially they could disapprove of the eight extraction alternatives at the same time.

[8] The reason why the punisher was charged by the total disapproval points and not just for the ones that corresponded to effective fishnet choices is that it was much simpler to explain and it enabled the subject to calculate the cost by himself. We consider that simplifying mechanisms is particularly important in a framed field environment like ours.

punished in the field (gossip, direct criticism, etc.) and evaluate its effects on the next periods' extractive decisions. We consider that punishing others socially may also have a social cost to the punisher but we were not particularly interested in studying it; we just intended to show that NMP was not for free for the punisher.

Punishment points for actual choices were added up and yellow, orange and red flags were assigned in accordance with the ranges shown in Table A.4 (see Appendix). It was not possible for someone to receive a red flag with just one subject disapproving his fishnet choice.

We employed a hybrid strategy method to implement this treatment. Punishment points were assigned after the subjects had been informed of the total number of nets used by the subgroup and therefore subjects could determine the average number of nets used by the others. It is a "hybrid" strategy method because individuals made decisions in two stages (and not as in the classical strategy method, where both decisions [extraction and punishment] are made at the same time). Brandts and Charness (2010) argue that following a strategy method instead of a direct punishment treatment can lead to lower disapproval among individuals. Also, Blount and Bazerman (1996) argue that individuals are less concerned with fairness when simultaneously choosing between two outcomes than when considering each outcome separately. For this reason, we chose a hybrid strategy method, one that is more similar to assigning punishment based on knowing the effective fishnet choices of each of the other members of the subgroup, but that still preserves anonymity. We discarded the alternative of disclosing actual individual levels of extraction in a random order because we considered there was a risk that anonymity would be violated.[9]

3.3 The structure of the experiment

Subjects were recruited during a survey that took place in March 2011. The aim of the survey was to gather data on socioeconomic characteristics and environmental perception among the resource users of artisanal fisher communities in Rocha's coastal lagoons. At the end of the questionnaire, the interviewee was asked whether he would be interested in participating in an activity where he could earn on average 2 daily wages (30 US dollars), depending on the decisions he would make. A week before the experiment we visited the communities where we delivered flyers in person to people from the five communities, and we made phone calls to those who had already been surveyed but could not be located while we visited the communities.

The experiment was conducted in two sessions in November 2011. Both sessions took place at La Paloma, a town in the province of Rocha, Uruguay. The communities that

[9] Keeping anonymity both in individuals' extraction decisions and in NMP was a priority. Indeed, as Anderies et al. (2011) point out, working with communities in field experiments requires developing this task with responsibility, because the game may not end when experimenters leave, and this may have spillover consequences in their daily life.

participated in each session were determined randomly (Table 1). Contrary to most framed field experiments, in this study subjects were transported from the place where they lived to the town where the experiment took place.[10] The aim of this design was to make subjects from different communities meet. This required that fishermen leave their community to attend the activity. It was particularly cumbersome to convince subjects to travel, and we believe it was the main reason why the number of participants was not as high as desired.

Table 1: Characteristics of the experimental sessions

In-group: "Groups and subgroups with individuals belonging to the same community".
Out-group: "Groups and subgroups with subjects belonging to two communities".
NMP: "Expressing disapproval of others' extraction levels. Those punished receive flags".

	Subjects		Treatments by period			
	Included in analysis	Discarded [a]	1-5	6-10	11-15	16-20
Session 1						
Laguna de Rocha	8	3	ingroup	ingroup punishment	outgroup	outgroup punishment
Valizas	8	3	ingroup	ingroup punishment	outgroup	outgroup punishment
Session 2						
El Puente	12		outgroup	outgroup punishment	ingroup	outgroup punishment
Puerto los Botes	8		outgroup	outgroup-punishment	ingroup	outgroup punishment
Barrio Parque	8		outgoup	outgroup punishment	ingroup	outgroup punishment
Total	44	6				

[a] During session 1 the subjects who turned up from Laguna de Rocha and Valizas were not multiples of four so three subjects from each community were selected randomly to play in subgroups of three and were reshuffled solely among the six all the periods. They were not considered in the analysis.

When subjects arrived at the venue, they drew a number from a bag (one bag per community). This number represented their identifier, and assigned each subject into a group of either eight or twelve members for each stage. Within these groups, subjects would play a CPR game in subgroups of four. The out-group treatment implied subgroups in which two subjects belonged to one community and two to the other.[11] In order to avoid repeated game type of behavior as much as possible, after each period subjects were reshuffled among all subjects in a group of eight or twelve. The subgroups they would play in the 20 periods were predetermined by the identifier number. It was common knowledge that the matching procedure between periods was random and had been determined by the initial draw of participants' identifier numbers. After each period, the experimenters

[10] Buses for each community were hired to pick up participants and transport them to the venue.

[11] In session 2, as there was one community in which there were twelve subjects (El Puente) and in the other two there were eight, during the out-group treatment, subgroups were composed of two subjects from el Puente and two from one of the other two communities or three from El Puente and one from the other community. In all cases the out-group treatment implied mixing just two communities.

indicated to the participants which subgroup of four they would play in the next period; at the end of the first 10 periods, participants were told in which group they would then play in (this implied a change in treatment from in-group to out-group or vice versa). During session 1, subjects played in an in-group treatment during the first stage, while in session 2 we reversed this order (see Table 1). This design enabled us to control for order effects.

Once in subgroups of four members, subjects were asked to sit with their backs facing each other so that they could not see the others' choices. Each group was conducted by a moderator who gave the instructions throughout the game, plus a monitor for every subgroup of four. This ensured that subjects did not interact during the game, and that an experimenter was always available to explain them how to use the material.

Subjects received a payoff table and an earnings sheet where they kept a record of their decisions and points gained. The payoff table summarized the pay-off consequences of all combinations of own nets used and the total number of nets used by the other three members of a subgroup (see Table A.2 in the Appendix). The exchange rate was set at 100 points for 0.62 US dollars. When looking at the payoff table, subjects had to make a decision as to how many nets to use (minimum one, maximum eight), which they wrote on a slip of paper and handed it in to the experimenter. Once the four subjects had written out their decisions, the total number of nets used by the subgroup was announced so that each subject could calculate the number of points they had earned and write that figure on their earnings sheet. The explanation of the game followed Cardenas (2003). The actual experiment began once the moderator had conducted three rehearsal periods and once all questions from participants had been clarified. All decisions were made privately and individually and only the total extraction by the four players was publicly announced.

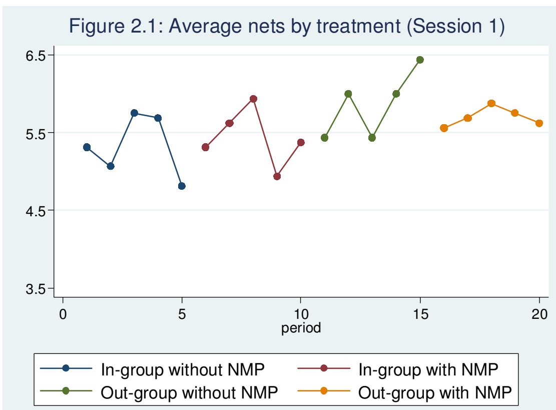
Before the punishing treatment started an example was provided. The example showed three subjects' disapproval cards: one punishing without any criteria, one punishing those who used many nets and one not punishing at all. The choice of nets and the disapproval points assigned were private information; the only public information was the flag received in case the subject was punished by the rest by more than one point. Subjects had to hold the flag so that others could see it during the next period of the game.

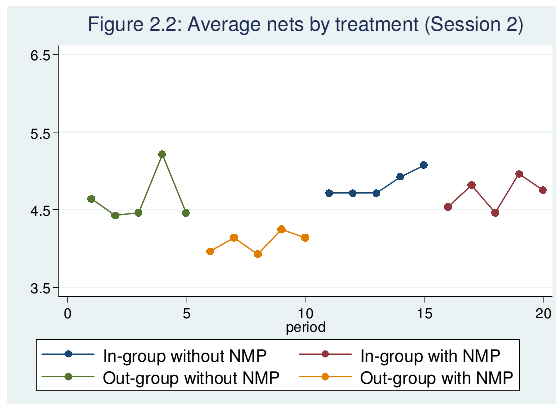
At the end of each experimental session we conducted a post-experiment survey which contained questions about reasons for disapproval, and feelings when being disapproved of. Each session of the experiment lasted about three hours and participants earned on average 30 US dollars (including a 5 US dollar show-up fee), a figure which amounts to 10% of a monthly minimum wage. [12]

[12] The experimental design excluding the in-group out-group treatment was tested with 36 undergraduate students.

4 Results

Figure 2 shows average extraction levels by period and treatment for session 1 and 2, respectively. At first glance, it suggests that the in-group/out-group treatment does not seem to induce significant changes in behavior when NMP is not available. Subjects' extraction levels in session 1 were higher during the out-group treatment without NMP but did not change substantially for subjects in session 2. The NMP treatment seems to have had a slight positive effect in terms of cooperation especially during the out-group treatment. It lowered average extraction levels in the second stage of session 1 and in both stages in session 2. It should also be noted that the three communities that participated in session 2 exhibited extraction levels significantly below those of the two communities that participated in session 1.





4.1 Testing treatment effects

Next we study players' extractive decisions in a dynamic analysis. This allows us to test the in-group out-group and NMP treatments. For this purpose, we estimate a dynamic model following a specification similar to that of Hayo and Vollan (2012), such that:

$$a_{it} = \alpha_i + \beta_1 \pi_{i;t-1} + \beta_2 \pi_{-i;t-1} + \beta_3 out-group\ w/NMP_{\ i} + \beta_4 in-group\ \ w/\ NMP_{\ i} \\ + \beta_5 in-group\ w/out\ NMP_{\ i} + \beta_6 stage_i + e_{it}$$

where:

$\boldsymbol{a_{it}}$ is *i*'s extraction level in period *t*, .$\boldsymbol{\pi_{i;t-1}}$ is individual *i's* payoff in a previous round. High payoffs in the previous round can be achieved either because there is cooperation (high group payoff and high individual payoff) or because of self-interested behavior (low group payoff and high individual payoff). Controlling for the group's payoff allows us to distinguish which of the two strategies is reinforced over time.

$\boldsymbol{\pi_{-i;t-1}}$ is the payoff of the rest of individual *i*'s subgroup (excluding individual *i*) in the previous round. Even if the game is a series of one-shot rounds and members of a subgroup change in every period, subjects may use information on the behavior of other subjects as a guide for future behavior. A negative relation between the group's payoffs in the previous period and the individual's extraction levels may suggest the existence of social preferences.

Treatments are tested in two ways. First, two dummy variables were included in the model: *in-group* that equals 1 if the players are under an in-group treatment (and 0 if they are playing the out-group treatment) and *NMP*, that equals 1 when the extraction decision is taken during a round that allows for NMP [rounds 6 to 10 and 16 to 20] and 0 otherwise). We also included a third dummy variable (*second stage*), which equals 1 for rounds 11 to 20. Second, we tested the interaction between treatments. For this purpose, three dummy variables were included: *out-group with NMP*, *in-group with NMP*, and *in-group without NMP* (*out-group without NMP is the base scenario*). Each of them equal 1 during the periods that they describe, and 0 otherwise. A fixed effects model was performed to control for individuals' time invariant characteristics. Time fixed effects were not included because they show high correlation with treatment variables (treatment dummy variables are time fixed effects).

Columns (1) to (6) in Table 2 show that while the in-group treatment has no effect on individuals' decisions, players chose lower extraction levels when playing during the NMP. Column (2) shows that the NMP treatment effect is significant independently of the additional variables included.

Table 2: Dynamic net decisions. Dependent variable: fishnets$_{it}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
earnings $_{i;t-1}$				0.004*	0.004*	0.004*		0.004*
				(0.002)	(0.002)	(0.002)		(0.002)
earnings $_{-i;t-1}$				-0.002*	-0.002*	-0.002*		-0.002*
				(0.001)	(0.001)	(0.001)		(0.001)
in-group	0.002		0.002	0.023		0.023		
	(0.139)		(0.139)	(0.128)		(0.128)		
NMP		-0.225*	-0.225*		-0.227**	-0.227**		
		(0.113)	(0.113)		(0.107)	(0.107)		
out-group with NMP							-0.414**	-0.401**
							(0.160)	(0.153)
in-group without NMP							-0.187	-0.159
							(0.199)	(0.187)
in-group with NMP							-0.223	-0.215
							(0.172)	(0.158)
second stage	0.402***	0.402**	0.402***	0.392***	0.383**	0.376***	0.402***	0.379***
	(0.139)	(0.161)	(0.139)	(0.130)	(0.149)	(0.132)	(0.139)	(0.131)
_cons	4.733***	4.847***	4.846***	5.165***	5.358***	5.352***	4.940***	5.387***
	(0.120)	(0.106)	(0.137)	(0.526)	(0.539)	(0.540)	(0.164)	(0.534)
Nº observations	880	880	880	836	836	836	880	836
Nº individuals	44	44	44	44	44	44	44	44
r2 within	0.019	0.025	0.025	0.031	0.037	0.037	0.029	0.041
r2 overall	0.009	0.012	0.012	0.180	0.180	0.180	0.014	0.171
r2 between	.	.	0.108	0.927	0.940	0.939	0.069	0.935

legend: *** p<0.01;** p<0.05; * p<0.1
robust standard errors in parenthesis

Heterogeneous treatment effects of NMP between in-group and out-group settings are shown in columns (7) and (8) of Table 2. On the one hand, it can be seen that the level of nets chosen under the out-group without NMP are not significantly different from the ones under the in-group, both with and without NMP. On the other hand, subjects under the out-group with NMP treatments extract lower levels than when the NMP is not allowed (the -0.4 coefficient amounts to 20% of a standard deviation in nets). Finally, the behavior of individuals under the in-group treatment is not significantly affected by the NMP treatment. The *second stage* dummy variable is positive and significant in all models. That is, subjects increase the average extraction level during the second stage, independently of the treatment they played first. The fact that cooperation decays throughout the game follows previous literature.

Regarding earnings in previous rounds, Models (4), (5), (6) and (8) shows that $\beta_1 > 0$, and $\beta_2 < 0$. This result is consistent with Hayo and Vollan (2012), and suggests that social preferences mechanisms are influencing players' decisions. As stated before, $\beta_1 > 0$, jointly with $\beta_2 < 0$ implies that individuals behave more cooperatively if their group in the past round performed well. This implies that their recent past experience influences their decisions, despite changing partners after each round.

4.2 Determinants of extraction decisions

In this section we analyze whether there are socio-demographic determinants of individual choices regarding extraction decisions. We do this for three variables of interest: number of nets chosen in the first period (columns 1 and 2), total number of nets chosen throughout the 20 periods (columns 3 to 5) and average nets (columns 6 to 8). Table 3 reports for each of these variables the general and reduced estimations.

Almost no individual-level economic and demographic variable seems to explain extraction choices, as Heinrich (2001) and Hayo and Vollan (2012) found. Wealth and age are the only observable individual determinants of choices which are significant. The magnitude of the wealth coefficient is worth noting: a one standard deviation increase in the wealth index increases the average choice of nets in 44% of a standard deviation. The wealth index was elaborated by means of factor analysis. The index considers different durable goods a household may own.[13] Cardenas (2003) also finds a positive relation between wealth and choices of extraction, and hypothesizes that low wealth status may reflect greater experience in managing a common pool resource. However, in our study this does not seem to be the case. Being a subject whose main activity is fishing is not related to extraction levels (see Table 3). Cardenas also provides an alternative explanation, which in our case can be understood if wealthier participants showed smaller marginal utilities from the cash earned in the experiment, thereby having less incentives to cooperate because the marginal

[13] The variables the index includes are the following: water heater, fridge, TV, radio, cable TV, DVD, washing machine, microwave, computer, Internet, phone, motorbike, car and horse.

value of potential gains is smaller than for the poorer participants. Hayo and Vollan (2012) report a positive coefficient on the upper middle and highest quartiles of income and also argue that high income might reveal a person's stronger preference for consumption, risk and competition.

The other significant determinant of fishnet choices is community membership. El Puente (the baseline in the regression) extracted significantly less than the other four communities. Also, the Wilcoxon-Mann-Whitney ranksum (WMW ranksum) tests reject median and mean extraction levels equality between places of residence, two-by-two, at 10% level of confidence (Table A.5 in Appendix). This hypothesis is not rejected only in the case between Barra de Valizas and Barrio Parque, with reference to average nets thrown, and between Laguna de Rocha and Barra de Valizas and Barrio Parque, with reference to average earnings during the experiment. However, median average earnings equality between the last two is rejected. These results, together with the non-significance of individual characteristics, strongly support the hypothesis that group level institutions or social norms influence individuals' behavior.

Table 3: Determinants of subjects' extraction decisions

	Nets first period		Total nets			Average nets		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Laguna de Rocha	0.84	1.79**	31.01**	40.33***	39.94***	1.55**	2.02***	2.00***
	(1.13)	(0.87)	(13.48)	(11.44)	(12.15)	(0.67)	(0.57)	(0.61)
Valizas	2.53**	1.17	52.92***	51.88***	50.16***	2.65***	2.59***	2.51***
	(1.08)	(0.87)	(12.93)	(11.60)	(12.08)	(0.65)	(0.58)	(0.60)
Botes	2.97**	1.42	25.05	24.22**	25.25**	1.25	1.21**	1.26**
	(1.26)	(0.87)	(15.06)	(11.15)	(11.73)	(0.75)	(0.56)	(0.59)
Barrio Parque	3.23**	1.42	38.35**	29.81**	31.23**	1.92**	1.49**	1.56**
	(1.29)	(0.87)	(15.35)	(11.49)	(12.13)	(0.77)	(0.57)	(0.61)
female	0.64		-5.46			-0.27		
	(0.86)		(10.28)			(0.51)		
age	-0.02		-1.07***	-0.53*	-0.58*	-0.05***	-0.03*	-0.03*
	(0.03)		(0.36)	(0.27)	(0.29)	(0.02)	(0.01)	(0.01)
years of schooling	-0.02		-3.50*			-0.17*		
	(0.17)		(1.98)			(0.1)		
drinkable water	-1.8		-6.39			-0.32		
	(1.08)		(12.83)			(0.64)		
electricity	-1.04		-17.02			-0.85		
	(1.06)		(12.61)			(0.63)		
wealth	0.49*		11.94***	8.04***		0.60***	0.40***	
	(0.29)		(3.4)	(2.90)		(0.17)	(0.14)	
per capita income (logs)	-0.99**		1.5			0.07		
	(0.48)		(5.7)			(0.29)		
fishing main activity	1.11		-2.29			-0.11		
	(0.77)		(9.17)			(0.46)		
perception[a]	-0.28		-3.85			-0.19		
	(0.77)		(9.17)			(0.46)		
trust[b]	-0.16		-13.11			-0.66		
	(1.21)		(14.4)			(0.72)		
second quartile (wealth)					7.00			0.35
					(11.13)			(0.56)
third quartile (wealth)					25.35**			1.27**
					(11.82)			(0.59)
fourth quartile (wealth)					27.02**			1.35**
					(12.72)			(0.64)
Constant	11.85**	3.83***	114.58**	72.59***	81.56***	5.73**	3.63***	4.08***
	(4.46)	(0.55)	(53.1)	(15.38)	(16.34)	(2.65)	(0.77)	(0.82)
Obs.	43	44	43	44	44	43	44	44
R-squared	0.35	0.12	0.61	0.46	0.45	0.61	0.46	0.45

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] Believes that preserving the environment in coastal lagoons is mainly a responsibility of the people rather than the government.
[b] Believes one can trust most people.

4.3 Types

In this section we classify subjects with reference to their actual cost of deviating from self-interested behavior, and their relationship with their subgroup partners into four categories: free riders, cooperators, conditional cooperators and others. The subject's actual cost of deviating from self-interested behavior is computed as the difference between the payoff the subject would have obtained in that period if he had extracted the maximum level (given the others' extraction choices) and the actual payoff he obtained. The classification was determined by analyzing individuals' behavior in the periods in which NMP was not available and subjects had already played one period and therefore had a reference point (periods 2-5 and 11-15).

Following the algorithm used by Kurzban and Houser (2005), we define a cost of deviating threshold equal to 18 points.[14] We consider a free rider to be a subject whose cost of deviating is below or equal to the 18 points threshold during all 9 periods. A cooperator is a subject whose cost of deviating is always above or equal to the 18 points threshold during these periods. In turn, a conditional cooperator is defined as a subject whose cost of deviating is both above and below the 18 points threshold and; who shows a positive slope in an ordinary-least-squares regression of own cost of deviating on the average cost of deviating from other members of his subgroup in a previous period. The slope measures the subject's responsiveness to others' behavior and could be interpreted as a proxy of λ mentioned in Section 2. Subjects that exhibit any other behavior in terms of their cost of deviating are classified as "other." The scatter plot of each subject's cost of deviation and others' average cost of deviation in the previous period are shown in Figure A.1 of the Appendix.

Table 4: Distribution and main characteristics by behavioral types

Type	Frequency	Average nets in 20 periods	Average total earnings (US)
Free rider	18%	6.9	28.31
Conditional cooperator	25%	4.9	28.22
Cooperator	20%	3.2	28.82
Other	36%	4.9	29.26

Table 4 reports the frequency of each type, average extraction levels and average earnings by type. There is a similar fraction of free riders and cooperators and a slight predominance of conditional cooperators. However, there is a substantial fraction of the subject pool which cannot be classified under any of these three types. It is worth noting that cooperators achieved higher earnings than free riders. This is due to a significant concentration of cooperators within some communities. Table A.6 in the Appendix shows that 50% of subjects coming from El Puente were classified as cooperators, enabling them to achieve greater earnings when playing together during the in-group treatment. In turn, Barrio Parque exhibits a high concentration of free riders, which lowered their earnings during the in-group treatment. Subjects classified as "Other" achieved the highest average earnings. In general, these subjects behaved very similarly to free riders. However, in some periods they increased their cost of deviating above the 18 point threshold, simultaneous to the rest of their partners. This allowed them to benefit from occasional synchronized cooperation and to thus obtain greater earnings.

[14] The threshold is set at 18 points because it is the median cost of deviating during the 9 periods.

4.4 Punishing behavior

In this section we analyze punishers' behavior. On average, 71% of subjects chose to punish in each period in which punishment was allowed. Disapproval was substantial throughout the game and was surprisingly quite high in the last period, even if subjects knew the experiment would be over after that period. Figure 3 presents average punishing points by period for the two sessions separately. It should be noted that in the out-group treatment, subjects were mixed among in-group and out-group members and did not know the extraction levels of each of them. Therefore, punishment could not be directly specifically to out-group members with certainty. Session1 exhibited higher levels of punishment during the out-group treatment, though this did not occur in session 2 in which the average disapproval levels are not significantly different in the out-group and in-group treatments. Considering the two sessions together, the amount of punishment is not significantly different in the out-group and in-group treatments.[15]

Figure 3.1: Average punishing points by period (session 1)

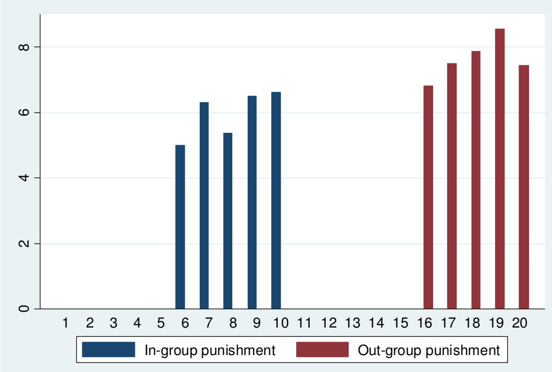


Figure 3.2: Average punishing points by period (session 2)

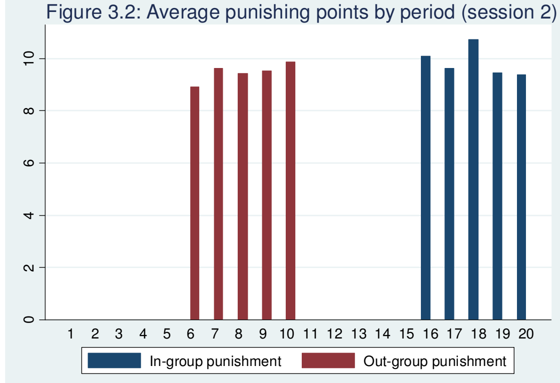
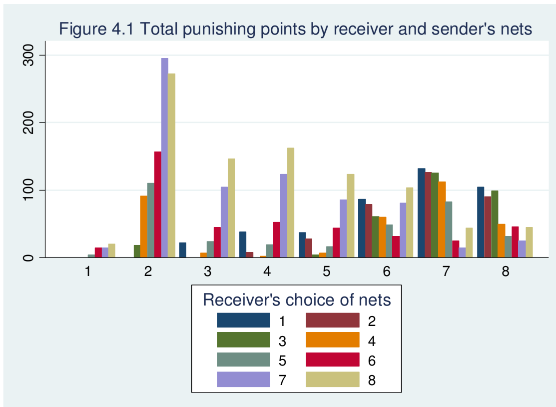


Figure 4.1 reports total disapproval points by the number of nets the subject chose to use in that period (horizontal axis) and to which choice of number of nets he decided to punish (bars). Those who use less than 6 nets disapprove of those who use more nets. This punishment of free riding could also be considered altruistic punishment as individuals incur material costs when punishing and reap no material benefits from punishing, because after punishing players are reshuffled before playing the next period. Also, we can observe antisocial punishment (punishment to cooperators): those who use 6 or more nets choose to disapprove of those who used fewer nets. Figure 4.2 shows per subject disapproval points instead of total points. Most subjects using 2 nets focused their disapproval on those throwing 6 nets and more, and only few subjects were spending a lot of disapproval points on those throwing 4 and 5 nets. There are only three subjects that used 6 or more nets and they spent a large number of disapproval points in lower extraction levels. This misdirected punishment is also observed by Falk et al. (2000), Masclet et al. (2003) and Gätcher and Herrmann (2011). Figures A.2.1 and A2.2 in the Appendix show average extraction levels in both sessions excluding sub-groups in which these three subjects participated. There, it can be observed that the effectiveness of the NMP treatment is greater than that observed in Figure 2.

15 A Wilcoxon-Mann-Whitney ranksum does not reject the equality between punishment directed during the out-group and in-group treatments for the two sessions together (*p-value*: 0.54).



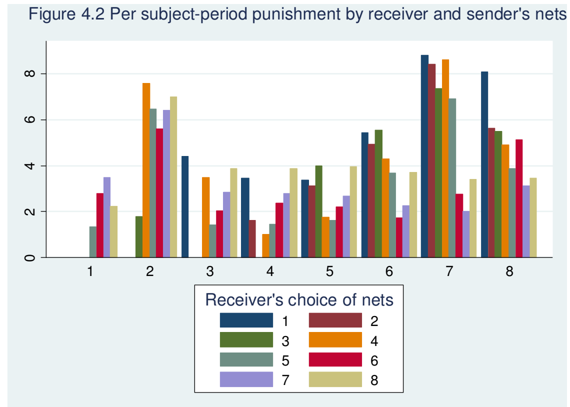


Figure 4 also indicates that there is some punishment from senders toward receivers using the same number of nets as themselves, especially when using a large number of nets. This could be interpreted as trying to discourage others from free riding while not sticking to the social norm in their actions (i.e., "do as I say and not as I do"). NMP was quite intense, when translating disapproval points into flags effectively received. We observe that on average there were 1.7 flags delivered per subgroup of 4 per period. Table 5 shows the distribution of flags received depending on whether the subject had chosen an extraction level below or above the subgroup's mean. Although the majority of the flags (60%) were awarded to subjects with extraction levels above the subgroups' mean, the remainder 40% were awarded to individuals with extraction levels below their subgroup's mean.

Table 5: Total flags awarded by round

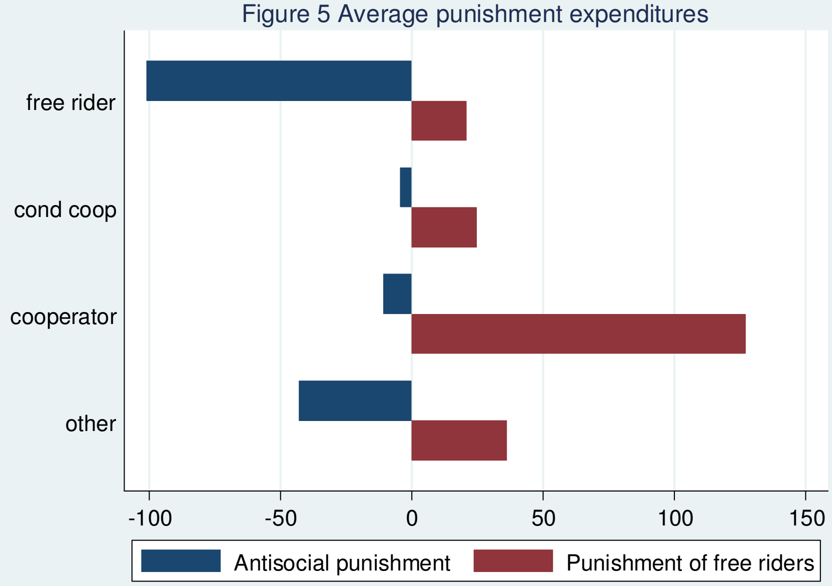
		Negative deviation max{0;āt-1-ai,t-1}				Positive deviation max{0;ai,t-1-āt-1}			
Period	Total	Total flags	Yellow	Orange	Red	Total flags	Yellow	Orange	Red
6	12	2	0	2	0	10	8	2	0
7	12	6	4	1	1	6	4	2	0
8	18	7	4	3	0	11	7	4	0
9	15	7	3	3	1	8	5	3	0
10	23	7	5	2	0	16	12	3	1
16	24	9	5	4	0	15	11	1	3
17	24	9	8	1	0	15	9	2	4
18	23	14	7	7	0	9	3	4	2
19	20	7	4	3	0	13	6	7	0
20	19	8	2	6	0	11	10	0	1
Total flags	190	76	42	32	2	114	75	28	11
%	100%	40.0%	22.1%	16.8%	1.1%	60.0%	39.5%	14.7%	5.8%

Cooperators were the ones who spent on average most points on disapproving others' behavior followed by free riders (see Table 6). As expected, cooperators were the least punished and free riders the most.

Table 6: Punishment sent and received by subject's type		
	Average cost of punishment	Average number of flags received
Free rider	122	5.4
Conditional Cooperator	29	4.2
Cooperator	138	3.0
Other	79	4.6

Following Herrmann et al. (2008), we consider punishment for extraction levels greater than one's own as punishment of free riding, and antisocial punishment to punishment directed to extraction levels equal to or smaller than one's own. Figure 5 shows average punishment expenditures following this categorization by subjects' types. Cooperators performed most punishment on free riders. In turn, free riders did most of the antisocial punishment although it was also observed to some extent in all the types. As with what was observed for overall punishment, antisocial punishment does not significantly differ between in-group and out-group treatments.

Figure 5 Average punishment expenditures



Among the post-experiment questions, we included one question regarding reasons for disapproving. Table 7 confirms that most disapproval points were directed to those who used many nets. However, from the three subjects who significantly disapproved of those who used few nets, two argued that they were disapproving those who threw few nets because they were missing out on chances to fish. They were the ones receiving the most punishment from others. The majority of subjects chose to disapprove of others' behavior because they were using too many nets (55%). It is worth noting that the average total cost of disapproval of the three subjects who disapproved of those who chose a low number of nets, was particularly high.

Table 7: Reasons for punishing				
Reasons for punishing	Average cost of disapproval	Average nets in punishment periods	Average number of flags received	Percentage of total subjects
Those who play different	144	6.8	6.5	5%
Those who threw many nets	84	4.3	3.7	55%
Did not disapprove	0	6.2	5.4	11%
Without any criteria	46	5.3	4.5	14%
Did not understand	21	3.3	3.0	5%
Those who threw few nets	376	6.4	6.7	7%
Part of the game	22	4.4	4.0	5%

Following Masclet et al. (2003) we estimated the following model:

$$P_{ik}^{t} = \beta_0 + \beta_1(\max\{0, a_i - a_k\}) + \beta_2(\max\{0, a_{av} - a_k\}) + \beta_3(\max\{0, a_k - a_i\}) + \beta_4(\max\{0, a_k - a_{av}\})$$

Where P_{ik}^{t} is the number of disapproval points that i assigns to k in round t, the coefficient β_1 is associated with positive deviations from the punisher's fishnet choice, that is, cases in which the punished chose fewer nets than the punisher, β_2 reflects the impact of positive deviations from the subgroup's average. In turn, β_3 reflects the relevance of negative deviations from the punisher's fishnet choice, that being situations in which the punished subject chose more fishnets than the punisher. Finally, β_4 is associated with the negative deviation from the subgroup's average. We included individual fixed effects to control for individuals' time invariant characteristics. We estimated the following model for each fishnet choice that could be punished. For instance, the first column in Table 8 reflects the determinants of punishing those subjects who chose 1 fishnet. As Table 8 shows, both positive (antisocial punishment) and negative (punishment of free riding) deviations from the punisher's fishnet choice are significant. But as Masclet et al. (2003) showed, there is an additional effect regarding deviations of the punished subject from the subgroup's average. [16]

[16] Estimates from a Tobit model point to the same conclusions but in that model, coefficients are slightly smaller in magnitude.

Table 8 Determinants of disapproval points directed to each of the fishnet options	1_net	2_nets	3_nets	4_nets	5_nets	6_nets	7_nets	8_nets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Positive deviation from i's own extraction ($\max\{0, nets_i - nets_k\}$)	0.79***	0.53***	0.89***	0.48***	0.67***	2.22***	1.46***	
	(0.07)	(0.07)	(0.10)	(0.10)	(0.12)	(0.15)	(0.42)	
Negative deviation from i's own extraction ($\max\{0, nets_k - nets_i\}$)				1.33***	0.35***	0.49***	0.33***	0.04
				(0.28)	(0.09)	(0.06)	(0.06)	(0.06)
Positive deviation from average ($\max\{0, nets_{av} - nets_k\}$)	0.13	0.31***	0.03	0.83***	0.70***	0.54	0.31	
	(0.09)	(0.10)	(0.15)	(0.16)	(0.20)	(0.52)	(1.91)	
Negative deviation from average ($\max\{0, nets_k - nets_{av}\}$)			-0.43	-0.16	0.32*	0.26***	0.29***	0.62***
			(0.62)	(0.32)	(0.16)	(0.09)	(0.08)	(0.08)
Constant	0.24***	0.28***	0.30***	0.42***	0.49***	0.38***	0.84***	0.98***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.08)	(0.10)
Observations	440	440	440	440	440	440	440	440
R-squared	0.55	0.47	0.44	0.32	0.24	0.51	0.29	0.30
Number of id_	44	44	44	44	44	44	44	44

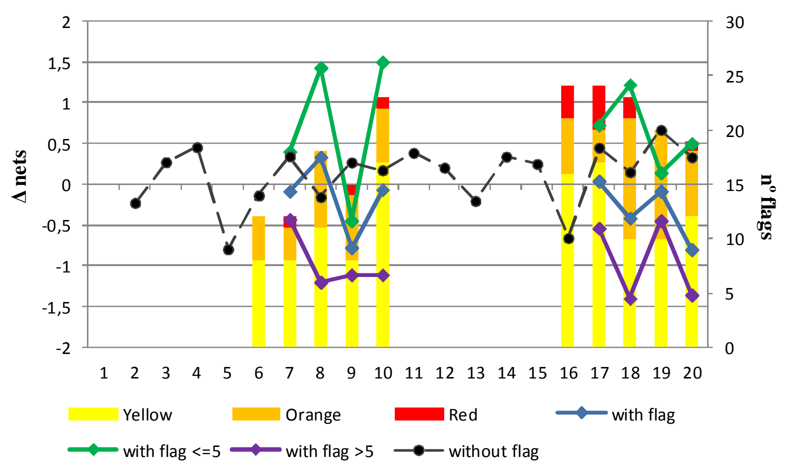
Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.5 Reaction to punishment

In this section we analyze whether punishment generated a change in behavior among those who were punished. At a first glance, the descriptive analysis suggests that flags produce variations in individuals' behavior. Figure 6 shows that individuals who received a flag in the previous period, on average, changed their behavior in the next period. However, at the individual level this is not always the case. One of the reasons for observing heterogeneity in terms of reaction to punishment is due to the fact that those who are punished are not only the ones who choose a high number of nets, but also those with a low number of nets. Figure 6 shows that while those who received a flag when throwing more than 5 nets diminish their choice in the next period, those who received a flag when throwing 5 or less nets increase the number of nets chosen in next period.

Figure 6: Total fishnets variations and total number of flags



Also, fishnets variations of those that receive flags seem to be greater, in absolute terms, for those who received a flag compared to those who did not receive one (Table 9).

Table 9: Changes in extraction

	n	Mean	s.d.	Min	Max
received a flag in t-1	148	-0.26	2.03	-7	5
received a flag when choosing nets >5 in t-1	81	-0.99	1.91	-4	5
received a flag when choosing nets <=5 in t-1	67	0.61	1.83	-7	2
did not receive a flag in t-1 (during NMP)	204	0.25	1.84	-7	7
did not receive a flag (periods 1 - 20)	688	0.07	1.92	-7	7

However, 33% of the individuals who received a flag did not change their behavior in the next round. This proportion is higher between those who decided to throw two (55%) and eight (48%) nets in the previous period (Table A.7 in Appendix). Decisions to throw two or eight nets are the modes of nets' distribution. A large number of subjects who chose these values decided not to change their choice, independently of what others think. Noussair et al. (2011) argue that subjects may not view the norm of cooperation as the norm that punishment should enforce, but that other norms such as "try to fish as much as possible" may be the prevailing ones. Therefore, some punished subjects may interpret punishment for using many nets as inappropriate and respond by raising the number of nets or maintaining their choice at the maximum number of nets.

Not all the flag colors produce the same reaction (see flag range in Table A.4 in Appendix). Subjects are more indifferent to yellow flags than to the others: 42% of the cases in which a subject received a yellow flag, he did not change his decision in the next period (Table A.8 in Appendix). When analyzing subjects' reaction in relation to how they said that they felt when receiving a flag in the post-experiment survey, those that declared indifference did not change their behavior after receiving a flag, or increased their decision in almost 70% of the cases (Table 10). Also, those who answered that they felt uncomfortable diminished their decision in the next period 52% of the times they received a flag (while in 28% of the cases, they maintained the decision they had made in the previous period). Table 10 also confirms that those subjects who considered they had been punished unfairly raised their extraction levels in the following period because they experienced anger.

Table 10 Nets variations and feelings (%)

	Feeling when receiving a flag - Total						Feeling when receiving a flag If nets in previous round <=5						Feeling when receiving a flag If nets in previous round >5					
nets variation	Unc.	Angry	Indif.	Fair	Other	Total	Unc.	Angry	Indif.	Fair	Other	Total	Unc.	Angry	Indif.	Fair	Other	Total
-	52	0	30.3	36.59	26.7	35.1	25	0	36	20.84	20	26.9	76.9	0	26.8	58.82	30	42.0
=	28	0	42.4	19.51	40	33.1	41.7	0	16	12.5	20	19.4	15.4	0	58.5	29.41	50	44.4
+	20	100	27.3	43.91	33.34	31.8	33.3	100	48	66.67	60	53.7	7.7	0	14.6	11.76	20	13.6
Total	100	100	100	100.01	100.01	100.01	100	100	100	100.01	100	100	100	0	100	99.99	100	100

Regarding behavior by type, when being punished for choosing more than 5 nets, conditional cooperators reacted by lowering extraction levels 50% of the time (as well as others), whereas free riders lowered extraction levels 26% of the time (Table A.9 in Appendix).[17] Instead, when being punished for throwing 5 or less nets, free riders reacted by increasing their extraction levels 100% of the time, conditional cooperators 65% and cooperators 38%, respectively.

The next step is to formally test for the behavior depicted above. We adapt the reaction function included in Masclet et al. (2003) and Noussair and Tucker (2005) and test whether player´s i decision changes from period t-1 to period t is a function of the punishment received in the previous period, and his extraction deviations from group average decisions:

$$a_i^t - a_i^{t-1} = \beta_0 + \beta_1 * YF_i^{t-1} + \beta_2 * OF_i^{t-1} + \beta_3 * RF_i^{t-1} + \beta_4 * (max\{0, a_i^{t-1} - \bar{a}^{t-1}\}) + \beta_5 * (max\{0, \bar{a}^{t-1} - a_i^{t-1}\})$$

[17] By definition cooperators never chose more than 5 nets.

Table 11: Net variations, flags and deviations from group's average in previous round

	Dependent variable: $Fishnets_{i,t} - fishnets_{i,t-1}$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Periods	**7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1-5 & 11-15**	**7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**
Positive deviation from average; $max\{0; a_{i,t-1} - \bar{a}_{t-1}\}$					-0.554***	-0.967***	-0.93***	-0.134	-0.684***	-0.967***	-0.156	-0.658***
					(0.119)	(0.179)	(0.17)	(0.456)	(0,173)	(0.163)	(0.448)	(0.159)
Negative deviation from average; $max\{0; \bar{a}_{t-1} - a_{i,t-1}\}$					0.917***	0.832***	0.84***	0.664***	0.068	0.826***	0.633***	0.117
					(0.152)	(0.144)	(0.15)	(0.156)	(0,511)	(0.144)	(0.159)	(0.491)
Flag in t	-0.61*						-0.26	-0.062	-0.159			
	(0.32)						(0.22)	(0.268)	(0,309)			
Yellow flag in t-1		-0.471	-0.033	0.004						-0.256	-0.010	-0.058
		(0.328)	(0.373)	(0.314)						(0.241)	(0.373)	(0.321)
Orange flag in t-1		-0.793*	-0.029	-0.861						-0.329	-0.221	-0.609
		(0.467)	(0.450)	(0.583)						(0.333)	(0.408)	(0.516)
Red flag in t-1		-1.105	2.802***	0.181						0.340	1.824	0.453
		(1.205)	(0.993)	(0.508)						(0.670)	(1.382)	(0.632)
_constant	0.29**	0.300**	0.566***	-0.702***	-0.178	0.132	0.21	-0.132	0.363	0.236	-0.107	0.315
	(0.14)	(0.138)	(0.088)	(0.194)	(0.163)	(0.172)	(0.19)	(0.205)	(0,345)	(0.178)	(0.215)	(0.331)
Nº observations	352	352	213	139	396	352	352	213	139	352	213	139
Nº individuals	44	44	39	32	44	44	44	39	32	44	39	32
r2 within	0.020	0.023	0.033	0.036	0.291	0.350	0.354	0.141	0.159	0.356	0.158	0.178
r2 overall	0.018	0.018	0.020	0.023	0.185	0.196	0.197	0.081	0.099	0.199	0.088	0.110
r2 between	0.013	0.001	0.001	0.002	0.021	0.038	0.039	0.000	0.036	0.040	0.000	0.022

legend: *** p<0.01;** p<0.05; * p<0.1
robust standard errors in parenthesis

Where YF_i^{t-1}, OF_i^{t-1}, and RF_i^{t-1}, are dummy variables that indicate if the individual received a yellow, orange or red flag, respectively, in a previous period, $max\{0, a_i^{t-1} - \bar{a}^{t-1}\}$ is a variable that indicates if the individual extracted more than his subgroup average in a previous period, and the deviation magnitude, while $max\{0, \bar{a}^{t-1} - a_i^{t-1}\}$ is the same but for negative deviations from the subgroup average in a previous period.[18] We test this model for periods where flag reaction could take place (periods 7 to 10 and 17 to 20) and separately for those that chose 5 or less nets and more than 5 nets, respectively. A control model during periods where reaction is not possible is also included in column (5) to compare conformity effects. We included individual fixed effects to control for non-observable factors that may affect individual decisions.

Model (1) in Table 11 suggests that being punished generates a downward adjustment in the following periods. When distinguishing by flag colors (Model 2), we observe that receiving an orange flag appears to have a small influence on diminishing individuals' nets choice, but its effect is diluted when splitting the sample between those people who receive a flag when they threw five or less nets, and more than five (Models 3 to 4). It is worth noting the large increase in fishnets choices in *t* when receiving a red flag, having thrown five or less nets in *t-1*. Subjects react strongly, in a non-cooperative way, when they feel they have been unfairly punished.

When conformity effects are allowed (how individuals deviate from the subgroup's average in the previous period), receiving a flag is not significant in any case any longer (Models (5) to (12)). Those who threw fewer nets than the subgroup's average in the previous period increase their decision in the next period, while those who threw more than average in previous period decrease their decision the next period. The presence of conformity effects is consistent with Masclet et al. (2003), and Hayo and Vollan (2012). As might be expected, the second mechanism does not take place if we look only at the reaction of those who received a flag when throwing five or less nets, while the first mechanism does not work when the reaction of those who threw more than five nets is studied. The magnitude of the conformity effect is larger during the NMP periods (especially positive deviations of the subjects' extraction relative to the subgroup's mean), which could indicate that there may be an additional impact of NMP increasing the convergence to the social norm. However, confidence intervals for these effects in periods with and without NMP overlap at the 95% confidence level. Table A.11 in the Appendix shows that interactions between deviations from the subgroup in previous periods and having received a flag are not significant, which confirms that the high significance of conformity effects do not seem to be picking the effect of being punished. Also, the conformity effects are not different when individuals are playing solely with people of their own community, relative to the out-group treatment (Table A.11 in Appendix).[19]

[18] We also include a specification in which instead of distinguishing different flag colors, there is just a single dummy variable that indicates that the subject received a flag in the previous period.

[19] We also estimated specifications that replicate those of Masclet et al. (2003) and Noussair and Tucker (2005), using the total number of punishment points received in the previous period instead of the flags

To sum up, as shown in section 4.1, the NMP treatment has an effect that reduces extraction levels, especially during the out-group treatment. However, when analyzing period-by-period variations in extraction decisions, individuals adjust their choice to the subgroup's average in the previous period rather than react to punishment. This conformity effect is present both when NMP is available and when it is not. Only those who are punished with a red flag and perceive that action as unfair appear to react by raising their extraction levels. This may be explained because they experience anger. However, this effect dilutes when taking conformity into account. The fact that receiving a flag does not have consequences on individuals' decisions can be explained because subjects who are sensitive to NMP lower their extraction levels in advance, to avoid being punished and experiencing shame. Indeed, they correctly anticipate that in order to reduce the probability of being punished the best they can do is lower their extraction levels. Subjects are aware that if they choose high extraction levels they are likely to be punished and if they choose to do so it is because the punishment does not generate a significant disutility. This explains why when individuals are punished they do not react to punishment (unless they did not expect it, as in the case of being punished by antisocial punishers).

5 Discussion

In this study we performed a framed field experiment to test the effectiveness of non-monetary punishment (NMP) in the context of a CPR game. We combined this treatment with an in-group/out-group treatment, letting fishermen from different communities play one stage of the experiment solely with members of their own community and the other stage mixed with another community.

Our findings suggest that NMP has an effect diminishing extraction levels only in the out-group treatment. Subjects derive more disutility from being punished when interacting with subjects who do not belong to their own community. Following the theoretical model in Section 2, this would imply that the σ_i coefficient (a measure of one´s susceptibility to social punishment) is greater when interacting in an out-group than in an in-group environment. In other words, subjects take the NMP institution more seriously during the out-group treatment. In a context in which individuals do not know each other (or hardly know each other) but are aware that there is a slight chance they might see each other again,

---

dummy variables. Results are consistent with the model shown before, but we prefer to keep the former adaptation, because it better reflects the information that individuals had when they made their decision. Also, Masclet et al. (2003) and Noussair and Tucker (2005) included overall deviations from the group's average in the previous period instead of positive and negative deviations when modeling decision changes. We also tested their specification, obtaining similar results, but we prefer to stick to our model, because as shown before, we would expect different reactions from those punished when throwing low number of nets than those punished when throwing high number of nets. The model above shows consistent results when changing each flag color dummy variable for a unique dummy variable, indicating that the individual received a flag of any color in the previous period (Table A.10 in the Appendix).

being publicly punished would provide the only information others have about oneself and in this sense it may be important to avoid being flagged in such a way. However, the NMP may not be perceived as intimidating when coming from workmates or neighbors. NMP may not matter either if it takes place in a context of complete strangers in which subjects know for sure they will not meet again. In other words, the relationship between the sensitivity to peer punishment in in-group/out-group contexts may be non-monotonic.

Previous literature regarding contributions in public good games finds that non-monetary punishment increases cooperation in a public good game, but its effect is smaller than that of monetary sanctions (Masclet et al., 2003), and it is more effective in increasing cooperation when combined with this kind of sanction (Noussair and Tucker, 2005). Our findings are consistent with these studies in pointing that non-monetary punishment, solely by affecting pro-social emotions, can enhance cooperation. However, its effects are not as strong as those of monetary punishment, which affects not only individuals' pro-social emotions but also their monetary payoffs.

The NMP's effectiveness is diminished by the fact that not only free riders but also cooperators are punished. Indeed, Beckenkamp and Ostmann (1999) and Masclet et al. (2003) report that if subjects perceive the sanctions as unfair, they can react by decreasing cooperation. The latter interpret punishment from non-cooperators as evidence of spiteful preferences. This misdirected punishment is also observed by Falk et al. (2000), and Gächter and Herrmann (2011). Herrmann et al. (2008) point out that one plausible explanation of antisocial punishment is that people might not accept punishment and therefore seek revenge. This is likely, as these subjects were being constantly punished by the rest. In fact, most punishment administered to free riders was performed by cooperators, while most of the antisocial punishment came from free riders. Alternatively, it could also be interpreted as features of their daily lives that subjects bring into the game (Cardenas and Ostrom, 2004). For instance, they may perceive that intensifying current fishing does not have any consequences on the availability of fish in the future (for instance, because they may believe that climate factors or other industries are more important determinants of fish availability). In the same line of thought, Casari and Luini (2009) and Noussair et al. (2011) argue that subjects may not view the norm of cooperation as the norm that punishment should enforce, as other norms such as "try to catch as many fish as possible" may be the prevailing ones. A fourth explanation could be that this behavior is a consequence of bounded rationality, related to cognitive limitations of the game on the part of some players, in line with Simon (1955). Janssen et al. (2010) argue that in a context in which participants can punish back but cannot discuss why they are sanctioned, receiving a sanction does not carry a clear message.

It is particularly interesting to note that subjects are willing to punish others while facing a monetary cost to themselves and may not necessarily expect that this punishment will determine an increase in cooperation. Even if the monetary cost of social punishment was low, subjects were reminded at every period that by socially punishing others they were themselves bearing a cost, as they had to subtract the total cost of punishment from their earnings in their balance sheet. Despite this fact, subjects chose to punish others during the whole experiment, including the last period when no change in others' behavior was possible. In fact, on average per period each subgroup awarded 1.7 disapproval flags to the members of that group. This result is in line with Fehr and Gachter (2000) findings regarding monetary punishment. Following Casari and Luini (2009), Fudenberg and Pathak (2010) and Noussair et al. (2011), we conclude that punishment is not necessarily applied instrumentally to increase cooperation and that subjects have preferences for punishing.

We do not find significant differences in punishing behavior between in-group and out-group treatments. This finding is in contrast to McLeish and Oxoby (2007) and Miguel and Gugerty (2005), who argue that subjects punish free riders more harshly in in-groups than out-groups. On the contrary, Chen and Li (2009) and Currarini and Mengel (2012) find that subjects are less likely to punish in-group members than out-group members.

Subjects do not adjust their period-by-period decisions as a reaction to punishment effectively received. They seem to correctly anticipate that the likelihood of being punished is increasing in extraction levels and those who would experience disutility by being punished reduce their extraction levels beforehand. Those who do not reduce extraction levels do not react to punishment because they are insensitive to it. Instead, those who were unexpectedly punished and who considered the punishment unfair, experienced anger and increased their extraction levels in the subsequent period.

We find strong conformity effects: individuals adjust their period-by-period decisions in order to converge with their peers' average in a previous period. These results are consistent with Velez et al. (2009) and Hayo and Vollan (2012). The results highlight the potential relevance of social comparisons as a form of non-pecuniary policy seeking changes in behavior (Ferraro and Price, 2011).

Contrary to what has been mostly documented in the literature, we do not find an in-group bias regarding cooperation. That is, individuals do not behave differently when interacting with subjects from their own community than when they are mixed with subjects from other communities, except for being more sensitive to NMP during the out-group treatment. Hewstone et al. (2002) argue that negative feelings toward out-group members tend to occur mostly in circumstances in which belonging to a group draws a strong sense of identity, and that this can be reduced as a consequence of the quantity and the quality of contact between groups. Fishermen from different communities do not interact during their daily life, but they are used to seeing each other while fishing as they tend to move from one lagoon to the other depending on fish availability. The non-relevance of the in-group favoritism may be explained because of the high mobility across lagoons. When surveyed, those fishermen who live by the shores of the coastal lagoons complained about others coming to fish there. But they usually also move to other places to fish, depending on fish availability. Therefore, they are used to seeing others fishing in their own place of residence, and even if they complain they know the same stands for themselves when they go to fish to another location. In other words, everybody acknowledges being an outsider at some point in time. For this reason, as regards to social preferences, granting exclusive access to a common pool resource to a certain community appears not to be a requisite from a resource conservation point of view.

Finally, community membership appears to have an influence over individuals' decisions, a finding not explained by observable socioeconomic factors. This may suggest that social norms regarding extraction levels differ among communities. The importance of community membership has been noted by Henrich et al. (2001) and Hayo and Vollan (2012). In our case it is quite striking to find differential behavior by community, as the communities we studied do not differ in terms of ethnicity or economic organization. Also, in line with other studies (Cardenas, 2003; Hayo and Vollan, 2012), we do find that cooperation is negatively correlated with wealth. This relationship should be studied more in depth, in order to disentangle the causal link between the two.

Overall, our results are consistent with the view that cooperation in a CPR dilemma is determined not only by repeated game behavior but also by social preferences. Subjects are willing to bear costs due to deviations from the self-interested equilibrium, even in one-shot interactions, as has been previously observed in public goods settings. However, previous interactions with other subjects have substantial influence on behavior, reflecting strong preferences for conformism. Individuals with social preferences limit their resource exploitation (cooperate) in response to the threat of punishment, but we do not find evidence of reactions to being effectively punished. We argue that the latter result is due to two reasons. First, subjects anticipate that the probability of being punished increases with their extraction level decision. Therefore, they reduce their extraction decision in advance, avoiding the experience of shame. Second, antisocial punishment was substantial and generated in some cases an increase in extraction among those being unfairly punished. Finally, our results suggest that the relevance of in-group favoritism in promoting cooperation due to social preferences may be overrated, and that for peer punishment to be effective it requires coordination, in order to prevent anti-social targeting and to enhance the social signal conveyed by the punishment.

Acknowledgments

We are particularly thankful to Sam Bowles, Andres Rius, Jordi Brandts, Francisco Alpizar and Fredik Carlsson for their invaluable advice. We are also very thankful for the comments received from Gabriele Camera, Luigi Luini, Jorge Maldonado, Shachar Kariv, Jeroen van der Bergh, Josep Lluís Raymond and Paolo Pin. We would also like to thank comments from the participants of the LACEEP's XII and XIII workshops and Javier Vitancurt, (DINAMA), and Graciela Fabiano (DINARA), for their help while collecting technical information. We are also grateful for the research funding provided by LACEEP and the contribution of Instituto de Economía (UdelaR).

This project would have never been possible to implement without the commitment and support of Hugo Bai, Gabriel Burdín, María Dolores Carnevale, Paula Carrasco, Marco Colafranceschi, Andres Dean, Inés Erro, Sandra García, Federico Méndez, Sybille Mertens, Cecilia Oreiro, Sylvina Porras, Lorena Repetto, Andrés Rius, Nicolás Schmidt and Andrea Vigorito during the experimental sessions. We are especially grateful to Cecilia Toledo, for her support and patience helping us coordinate the field work. We are also grateful to La Paloma town hall, for the facilities provided. Finally, we would like to thank the fishermen from the five communities who participated in the experiment. All errors are our own.

References

Akerlof, G., and Kranton, R. (2000) "Economics and identity". *Quarterly Journal of Economics*, 115(3): 715–53.

Akerlof, G., and Kranton, R. (2005) "Identity and the economics of organizations". *Journal of Economic Perspectives*, 19(1): 9–32.

Anderies, J.M., et al. (2011) "The challenge of understanding decisions in experimental studies of common pool resource governance" *Ecological Economics*, 70, pp. 1571 – 1579.

Baland, J.M., Platteau, J.P. (1996) "Halting degradation of natural resources. is there a role for rural communities?". *Clarendon Press Oxford*.

Baland, J.M., Bardhan, P. and Bowles, S. (2007) *"Inequality, cooperation and environmental sustainability"*. Russell Sage Foundation and Princeton University Press.

Bandiera, O.; Barankay, I. and Rasul, I. (2005). "Social Preferences and the Response to Incentives: Evidence from Personnel Data". *Quarterly Journal of Economics*, 120(3): 917–62.

Barr A. (2001) "Social Dilemmas and Shame-based Sanctions: Experimental results from rural Zimbabwe".Working Paper 2001.11, Centre for the Study of African Economies University of Oxford.

Beckenkamp, M., and Ostman, A. (1999) "Missing the target? sanctioning as an ambiguous structural solution". in Foddy, M., Smithson, M., Scheneider, S. and Hogg, M. (eds.) *"Resolving social dilemmas: dynamic, structural, and intergroup aspects"*. London: Psychology Press, 165-180.

Bernhard H., Fehr, E. and Fischbacher, U. (2006) "Group affiliation and altruistic norm enforcement". *American Economic Review*, 96(2): 217–21.

Blount, S. and Bazerman, M. (1996), "The inconsistent evaluation of absolute versus comparative payoffs in labor supply and bargaining". *Journal of Economic Behavior and Organization*, 30, 227-240.

Bowles, S. (2004) *"Microeconomics: behavior, institutions and evolution"*. Russell Sage Foundation.

Bowles, S. and Gintis, H. (2002) "Social capital and community governance". *The Economic Journal* Vol. 112 N. 483: 419-436.

Bowles, S. and Gintis, H. (2011) *"A cooperative species"*. Princeton University Press.

Brandts J. and Charness, G. (2010) "the strategy versus the direct-response method: a survey of experimental comparisons".

Casari, M. and Luini, L. (2009) "Cooperation under alternative punishment institutions: An experiment". *Journal of Economic Behavior & Organization*, 71: 273-282.

Cardenas, J.C. (2003) "Real wealth and experimental cooperation: experiments in the field lab". *Journal of Development Economics* 70 (2003) 263– 289.

Cardenas, J.C. (2004) "Norms from outside and from inside: an experimental analysis on the governance of local ecosystems." *Forest Policy and Economics* (6):229-241Chen , Y and Xin S. (2009) "Group identity and social preferences". *American Economic Review*, 99(1): 431–57.

Denant-Boemont, L., Masclet, D. and Noussair, C. (2011) "Announcement, observation and honesty in the voluntary contributions game". *Pacific Economic Review*, Vol. 16, Nº 2, pp. 207 – 228.

Dugar, S. (2010) "Nonmonetary sanctions and rewards in an experimental coordination game" Journal of Economic Behavior & Organization 73: 377 – 386.

Falk, A., Fehr, E. and Fischbacher, U. (2000) "Informal sanctions". Working paper, University of Zurich.

Feeny, D., Berkes, F., McCay, B. and Acheson, J. (1990) "The tragedy of the commons, twenty-two years laters". *Human Ecology*, Vol. 18, Nº 1.

Fehr, E. and Gächter, S. (2000) "Cooperation and punishment in public good experiments". *American Economic Review*, Vol. 90, Nº 4, pp. 980 – 994.

Ferraro, P and Price, M. (2011) "Using non-pecuniary strategies to influence behavior: evidence from a large scale field experiment". NBER working paper series N 17189.

Fudenberg, D. and Pathak, P. (2010) "Unobserved punishment supports cooperation". *Journal of Public Economics* 94: 78-86.

Gächter, S. and Fehr, E. (1999) "Collective action as a social exchange". *Journal of Economic Behavior & Organization*, Vol. 39, pp. 341–369.

Gächter, S. and Herrmann, B. (2011) "The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia". *European Economic Review*, 55, pp. 193 – 210.

Goette, L., Huffman, D. and Meier, S. (2006) "The impact of group membership on cooperation and norm enforcement: evidence using random assignment to real social groups". IZA Discussion Paper, Nº 2020.

Hayo B. and Vollan, B. (2012) "Group interaction, heterogeneity, rules, and cooperative behaviour: Evidence from a common-pool resource experiment in South Africa and Namibia". *Journal of Economic Behavior and Organization* 81: 9-28.

Hardin, G. (1968) "The tragedy of commons". in Stavins (2000) *"Economics of the environment. Selected Readings"*, W. W. Norton.

Harrison, G. and List, J. (2004) "Field experiments". *Journal of Economic Literature*, Vol. 42, Nº 4. (Dec., 2004), pp. 1009-1055.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. and McElreath, R. (2001) "In search of homo economicus: behavioral experiments in 15 small-scale societies". *American Economic Review*, Vol. 91, Nº 2, pp. 73 – 78.

Herrmann, B; Thoni, C.; and Gachter, S. (2008) "Antisocial punishment across societies". *Science* 319 (5868), 1345-6.

Hewstone, M., Rubin, M. and Willis, H. (2002) "Intergroup bias". *Annual Review of Psychology*, Vol. 53, pp. 575 – 604.

Janssen, M., Holahan R., Lee, A. and Ostrom, E. (2010) "Lab experiments for the study of socio-ecological systems". *Science* 328: 613-617.

Kurzban, R. and Houser D. (2005) "Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations". *Proceedings of the National Academy of the United States of America*, 102(5): 1803-1807.

López M.; Murphy, J.; Spraggon, J. and Stranlund, J. (2012) "Comparing the effectiveness of regulation and pro-social emotions to enhance cooperation: experimental evidence from fishing communities in Colombia". *Economic Inquiry*, Vol. 50 (1): 131-142.

Masclet, D., Noussair, C., Tucker, S. and Villeval, M.C. (2003) "Monetary and nonmonetary punishment in the voluntary contributions mechanism". *American Economic Review*, Vol. 93, Nº 1, pp. 366 – 380.

Miguel, E. and Gugerty, M.K. (2005) "Ethnic diversity, social sanctions and public goods in Kenya". *Journal of Public Economics*, Nº 89, pp. 2325 – 2368.

Noussair, C. and Tucker, S. (2005) "Combining monetary and social sanctions to promote cooperation". *Economic Inquiry* Vol. 43, No. 3 649-660.

Noussair, C. and Tucker, S. (2007) "Public observability of decisions and voluntary contributions in a multiperiod context". *Public Finance Review* Vol. 35 Number 2 176-198.

Noussair, C., van Soest, D. and Stoop, J. (2011) "Punishment, reward, and cooperation in a framed field experiment". Munich Personal RePEc Archive.

Ostrom, E. (1990) *"Governing the Commons: The Evolution of Institutions for Collective Action*"- New York: Cambridge University Press.

Ostrom, E., Burger, J., Field, C., Norgaard, R., and Policansky, D. (1999) "Revisiting the commons: local lessons, global challenges". *Science* Nº 284, 278 - 282

Ostrom, E. (2000) "Collective action and the evolution of social norms". *The Journal of Economic Perspectives*, Vol. 14, Nº 3.

PROBIDES (2002) "Proyecto de Desarrollo de la Areas Protegidas Lagunas de Garzon y Rocha". Documentos de Trabajo Nº 44.

Rege, M. and Telle K., (2004) "The impact of social approval and framing on cooperation in public good situations". *Journal of Public Economics* 88: 1625– 1644.

Rodríguez-Gallego, L., Santos, C., Amado S., Gorfinkel, D., González, M.N., Gómez, J., Neme, C., Tommasino, H. and Conde D. (2008) "Costos y beneficios socioeconómicos y ambientales del uso actual de la Laguna de Rocha y su Cuenca. Insumos para la Gestión Integrada de un Área Protegida Costera". Informe Programa de Desarrollo Tecnológico, UdelaR.

Ruffle, B. and Sosis, R. (2006) "Cooperation and the in-group-out-group bias: A field test on Israeli kibbutz members and city residents". *Journal of Economic Behavior & Organization*, Vol. 60, pp 147–163.

Tanaka, T.; Camerer, C. and Nguyen, Q. (2008) "Status, ethnicity, and wealth in vietnam: evidence from experimental games". Unpublished.

Thompson, D. (2007) "Economía e identidad de pescadores laguna de Rocha". Dirección de Cultura, Ministerio de Educación y Cultura.

Trivers, R. (1971) "The evolution of reciprocal altruism". *Quarterly Review of Biology*, 46: 35-57.

Simon, H.A. (1955) "A behavioral model of rational choice". *The Quarterly Journal of Economics*, Vol. 69, Nº. 1, pp. 99-118.

Van Soest, D. and Vyrastekova, J. (2006) "Peer enforcement in CPR experiments: the relative effectiveness of sanctions and transfer rewards, and the role of behavioural types", List, J. (ed.) *"Using Experimental Methods in Environmental and Resource Economics*". Edward Elgar.

Velez, M.A., Stranlund, J.K., and Murphy, J.J. (2009) "What motivates common pool resource users? Experimental evidence from the field". *Journal of Economic Behavior & Organization*, Vol. 70, pp. 485–497.

Appendix

Table A.1: Mean socioeconomic characteristics by community					
Community	**Years of schooling**	**Electricity**	**Wealth**	**Per capita income**	**Fishing main activity**
LR - Barra	6.0	0.13	1.85	2987	0.75
Valizas Puente	6.7	0.75	3.06	3506	0.67
Barra de Valizas	7.6	0.38	1.68	7463	0.63
Puerto los botes	6.0	1.00	2.52	4913	1.00
Barrio Parque	8.0	1.00	4.32	6409	0.38

Table A.2 Payoff table

	My fishnets								
Others' total	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Others' average**
3	354	360	366	372	378	384	390	396	1
4	342	348	354	360	366	372	378	384	1
5	330	336	342	348	354	360	366	372	2
6	318	324	330	336	342	348	354	360	2
7	306	312	318	324	330	336	342	348	2
8	294	300	306	312	318	324	330	336	3
9	282	288	294	300	306	312	318	324	3
10	270	276	282	288	294	300	306	312	3
11	258	264	270	276	282	288	294	300	4
12	246	252	258	264	270	276	282	288	4
13	234	240	246	252	258	264	270	276	4
14	222	228	234	240	246	252	258	264	5
15	210	216	222	228	234	240	246	252	5
16	198	204	210	216	222	228	234	240	5
17	186	192	198	204	210	216	222	228	6
18	174	180	186	192	198	204	210	216	6
19	162	168	174	180	186	192	198	204	6
20	150	156	162	168	174	180	186	192	7
21	138	144	150	156	162	168	174	180	7
22	126	132	138	144	150	156	162	168	7
23	114	120	126	132	138	144	150	156	8
24	102	108	114	120	126	132	138	144	8

Table A.3: Punishment card	
If the other throws:	I disapprove (0 to 10 points)
1 net	
2 nets	
3 nets	
4 nets	
5 nets	
6 nets	
7 nets	
8 nets	
Total	

Table A.4: Flag range	
Flag	Total punishment points received
Yellow	2 - 5
Orange	6 - 10
Red	11 - 30

Table A.5: Wilcoxon-Mann-Whitney ranksum and median equality average nets tests by place of residence

		Mean				
		LR - Barra	Valizas Puente	Barra de Valizas	Puerto Los Botes	Barrio Parque
Median	Laguna de Rocha	x	7.364	-1.764	3.891	-1.711
	p	x	*0*	*0.0777*	*0.0001*	*0.0871*
	El Puente	59.0262	x	-8.998	-4.869	-8.296
	p	*0*	x	*0*	*0*	*0*
	Barra de Valizas	2.6313	57.1846	x	5.714	0.037
	p	*0.105*	*0*	x	*0*	*0.9706*
	Puerto los Botes	4.564	14.7255	9.1414	x	-5.142
	p	*0.033*	*0*	*0.002*	x	*0*
	Barrio Parque	3.4133	59.0262	0.0514	9.8246	x
	p	*0.065*	*0*	*0.821*	*0.002*	x

Table A.6: Distribution of type by community

	LR	El Puente	Valizas	Los Botes	Barrio Parque	Total
Free rider	25%	0%	25%	13%	38%	7
Conditional cooperator	50%	17%	25%	25%	13%	12
Cooperator	0%	50%	25%	0%	13%	9
Other	25%	33%	25%	63%	38%	16
Total	100% 8	100% 12	100% 8	100% 8	100% 8	44

Figure A.1 Own current cost and others' lagged cost of deviating by subject

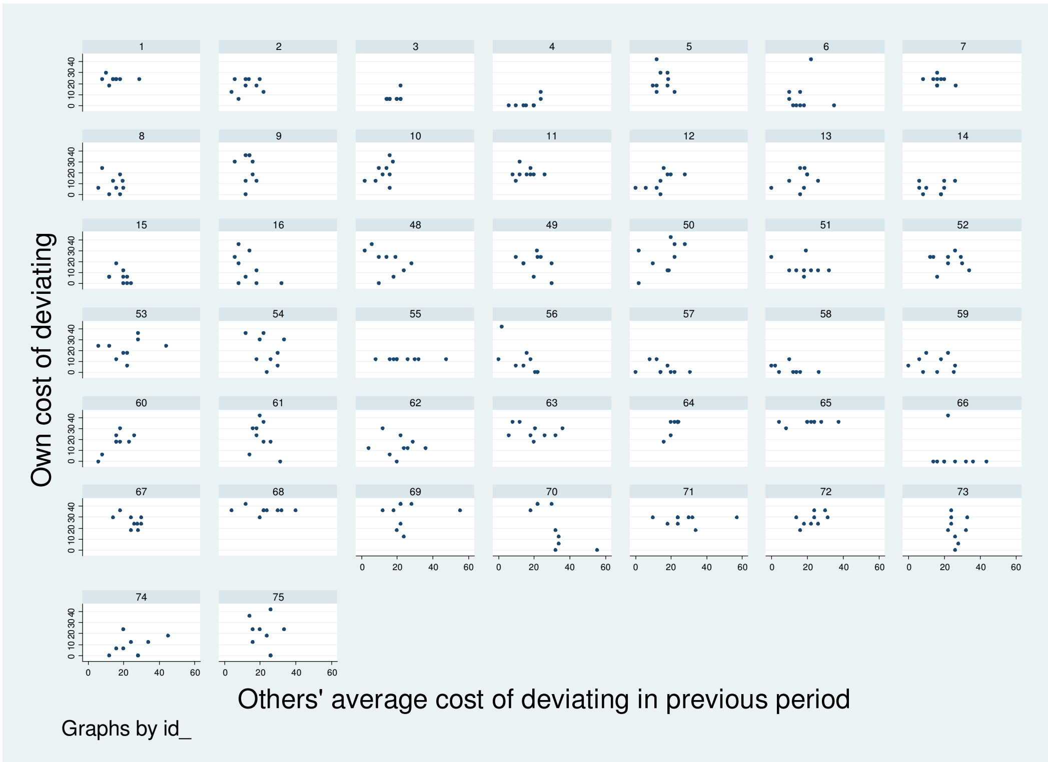


Table A.7: Nets variations and nets in previous round (in %)																			
	Total									If nets in previous round <=5						If nets in previous round >5			
	nets in previous round									nets in previous round						nets in previous round			
nets variation	1	2	3	4	5	6	7	8	Total	1	2	3	4	5	Total	6	7	8	Total
-7								3.23	0.68									3.23	1.23
-6							4.55	3.23	1.35								4.55	3.23	2.47
-5						7.14	4.55	3.23	2.7							7.14	4.55	3.23	4.94
-4					5	0	4.55	3.23	2.03					5	1.49	0	4.55	3.23	2.47
-3				9.52	5	7.14	13.64	6.45	6.76				9.52	5	4.48	7.14	13.64	6.45	8.64
-2			0	9.52	0	7.14	4.55	12.9	6.08			0	9.52	0	2.99	7.14	4.55	12.9	8.64
-1		18.18	9.09	9.52	35	14.29	4.55	19.35	15.54		18.18	9.09	9.52	35	17.91	14.29	4.55	19.35	13.58
0	25	54.55	18.18	9.52	10	50	31.82	48.39	33.11	25	54.55	18.18	9.52	10	19.4	50	31.82	48.39	44.44
1	0	0	54.55	28.57	25	7.14	31.82		17.57	0	0	54.55	28.57	25	25.37	7.14	31.82		11.11
2	25	18.18	18.18	14.29	15	7.14			8.78	25	18.18	18.18	14.29	15	16.42	7.14			2.47
3	0	0	0	4.76	5				1.35	0	0	0	4.76	5	2.99				
4	25	9.09	0	14.29					3.38	25	9.09	0	14.29		7.46				
5	25	0	0						0.68	25	0	0			1.49		.		
6	0	0							0	0					0				
7	0								0	0					0				
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Table A.8: Net variations and flag colour in previous round (%)

	Total			If nets in previous round <=5			If nets in previous round >5		
nets variation	**Yellow**	**Orange**	**Red**	**Yellow**	**Orange**	**Red**	**Yellow**	**Orange**	**Red**
-	28.4	44.9	45.5	25.0	32.1	0.0	30.8	61.9	62.5
=	42.1	20.4	18.2	25.0	14.3	0.0	53.9	28.6	25.0
+	29.6	34.7	36.4	50.0	53.6	100.0	15.4	9.5	12.5
Total	100	100	100	100	100	100	100	100	100

Table A9: net variations by behavioral type (%)

	Type - Total					Type - If nets in previous round <=5					Type - If nets in previous round >5				
nets variation	**free rider**	**cond. coop.**	**coop**	**other**	**Total**	**free rider**	**cond. coop.**	**coop**	**other**	**Total**	**free rider**	**cond. coop.**	**coop**	**other**	**Total**
-	25.01	47.1	19.1	42.4	35.1	0.0	35	19.1	28.0	27	25.8	50	0.0	53	42.0
=	59.4	17	43	25.4	33.1	0.0	0	42.9	16.0	19.4	61.3	38	0.0	32	44.4
+	15.6	41.7	38.1	32.2	31.8	100.0	65	38.1	56.0	53.7	12.9	13	0.0	15	13.6
Total	100	100	100	100	100	100	100	100	100	100	100	100	0	100	100

Table A10 Reaction to punishment other specifications.

	Masclet et al. (2003) & Noussair et al. (2011) model for punishment points received									Masclet et al. (2003) & Noussair et al. (2011) model controlling for flags received				
	Dependent variable: $fishnets_t - fishnets_{t-1}$													
Sample	**1-20**	**1-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1 - 20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**	**1-20**	**1-20**	**7-10 & 17-20**	**nets <=5 & 7-10 & 17-20**	**nets >5 & 7-10 & 17-20**
Punishment points in *t-1*	-0.054*	-0.003	-0.010	0.005	-0.046	-0.011	-0.012	-0.011	-0.054					
	(0.032)	(0.030)	(0.041)	(0.039)	(0.055)	(0.028)	(0.042)	(0.037)	(0.056)					
Flag in *t-1*										-0.408**	-0.118	-0.279	0.044	-0.300
										(0.191)	(0.173)	(0.223)	(0.271)	(0.316)
Deviation from average ($n^{t-1}_i - n^{t-1}_{av}$)		-0.781***	-0.910***	-0.655***	-0.682***						-0.778***	-0.903***	-0.655***	-0.711***
		(0.066)	(0.108)	(0.151)	(0.147)						(0.068)	(0.109)	(0.148)	(0.154)
Positive deviation from average ($n^{t-1}_i - n^{t-1}_{av}$)						-0.666***	-0.947***	-0.119	-0.663***					
						(0.098)	(0.166)	(0.450)	(0.165)					
Negative deviation from average ($n^{t-1}_{av} - n^{t-1}_i$)						0.849***	0.841***	0.669***	0.104					
						(0.109)	(0.148)	(0.162)	(0.490)					
_cons	0.076*	0.014	0.063	-0.081	0.402	-0.103	0.144	-0.135	0.418	0.093**	0.034	0.154	-0.085	0.464
	(0.039)	(0.037)	(0.108)	(0.138)	(0.288)	(0.118)	(0.188)	(0.197)	(0.343)	(0.039)	(0.035)	(0.094)	(0.150)	(0.338)
N	836	836	352	213	139	836	352	213	139	836	836	352	213	139
N_g	44	44	44	39	32	44	44	39	32	44	44	44	39	32
r2_w	0.006	0.306	0.370	0.172	0.205	0.288	0.350	0.142	0.165	0.007	0.306	0.374	0.172	0.206
r2_o	0.004	0.185	0.209	0.084	0.111	0.173	0.196	0.081	0.097	0.006	0.185	0.210	0.084	0.113
r2_b	0.188	0.125	0.031	0.000	0.041	0.126	0.037	0.000	0.042	0.071	0.126	0.032	0.000	0.044

legend: *** p<0.01;** p<0.05; * p<0.1

standard errors in parenthesis

Table A.11 Specifications including interactions		
Dependent variable: fishnets$_t$ -fishnets$_{t-1}$		
Sample	**7-10 & 17-20**	**7-10 & 17-20**
Positive deviation from average ($n^{t-1}_{i}-n^{t-1}_{av}$)	-0.952***	-0.94***
	(0.183)	(0.20)
Negative deviation from average ($n^{t-1}_{av}-n^{t-1}_{i}$)	0.876***	0.699***
	(0.166)	(0.169)
Positive deviation from average ($n^{t-1}_{i}-n^{t-1}_{av}$)*flag$_{t-1}$	-0.010	
	(0.204)	
Negative deviation from average ($n^{t-1}_{av}-n^{t-1}_{i}$)*flagt-1	-0.095	
	(0.168)	
Positive deviation from average ($n^{t-1}_{i}-n^{t-1}_{av}$)*outgroup		-0.08
		(0.17)
Negative deviation from average ($n^{t-1}_{av}-n^{t-1}_{i}$)*outgroup		-0.24
		(0.16)
_cons	0.120	0.14
	(0.165)	(0.17)
N	352	352
N_g	44	44
r2_w	0.351	0.356
r2_o	0.197	0.201
r2_b	0.038	0.040
legend: *** p<0.01;** p<0.05; * p<0.1		
standard errors in parenthesis		

Figure A.2 Average extraction levels by treatment and session excluding subgroups where three subjects performed substantial antisocial punishment