The perils of peer punishment: evidence from a common pool resource framed field experiment

Gioia de Melo

Universidad de la República, Uruguay

University of Siena, Italy

gioiademelo@iecon.ccee.edu.uy

Matías Piaggio

Universidad de la República, Uruguay

Universitat Autònoma de Barcelona, Spain

mpiaggio@iecon.ccee.edu.uy

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, MN, July 27-29, 2014.

Abstract

We provide experimental evidence on the effects of non-monetary punishment by peers among communities of Uruguayan fishers exploiting a common pool resource (CPR). We combined this treatment with an in-group (groups from a single community) / mixed group (groups composed of fishers from different communities) treatment. Our aim is to compare the effects of non-monetary sanctions in a context in which individuals exploiting a CPR belong to different communities relative to the case in which only individuals from the same community are allowed to exploit the resource. We find that mixed groups—unlike in-groups—reduce their exploitation of the resource in response to the threat of punishment. We do not find any differences in behavior between in-groups and mixed groups when the possibility of being punished is not available. The effectiveness of non-monetary punishment is reduced because cooperation was not perceived as the unique social norm. In such cases there is substantial antisocial punishment, which leads to increased extraction of the CPR by those who are unfairly punished. These findings indicate that effective peer punishment requires coordination to prevent antisocial targeting and to clarify the social signal conveyed by punishment.

1. Introduction

The exploitation of a common pool resource (CPR) poses a typical social dilemma. Hardin (1968) proposes the establishment of either private or state property rights as a solution to avoid the so-called tragedy of the commons. However, since informational asymmetries often vitiate the attempts of government regulations or market contracts to prevent overexploitation, communal property regimes have become an attractive alternative for the conservation and sustainable use of CPRs. It has been argued in many studies that, by enforcing social norms, communal property can fill the gaps of incomplete contracts (Feeny et al., 1990; Ostrom, 1990; Baland and Platteau, 1996; Ostrom et al., 1999; Ostrom, 2000; Bowles and Gintis, 2002).

Although much research has explored the determinants of successful communal property regimes, the issue is far from settled.[1] In this study, we evaluate whether nonmonetary punishment (NMP) is effective in promoting, via social preferences, cooperation in a CPR dilemma.[2] Nonmonetary punishment could prove to be a viable approach to community management because it is in such a setting that informal sanctions are typically applied. We seek to compare the effectiveness of NMP (concerning a CPR) when the exploiting individuals belong to different communities with its effectiveness when such individuals all belong to the same community. Different kinds of institutions may be effective depending on the context in which a CPR is exploited (Ostrom et al., 1999). Our motivation for studying this problem is reflected in two questions. First, given a particular structure of property rights (i.e., how many communities are granted the rights to exploit a CPR), how effectively do nonmonetary sanctions promote cooperation? Second, what are the advantages of different property right schemes? For example, a finding that social sanctions are more effective within than among communities would argue for CPR management rights being granted to one community rather than several.

Punishment is often viewed not only as a way to incentivize desired behavior but also as a "moral lesson" in condemning antisocial behavior (Bowles and Polanía-Reyes, 2012). We intend to test whether or not, in the absence of monetary incentives, moral lessons can effectively guide behavior toward socially beneficial ends. The prosocial emotions of an individual being punished are better identified by nonmonetary punishment than when she must endure monetary (costly) punishment. Findings in the literature on the effectiveness of costly and noncostly punishment are mixed. Whereas van Soest and Vyrastekova (2006) report that costly punishment effectively increases cooperation in a CPR dilemma, Noussair et al. (2011) observe no significant changes in cooperation. Janssen et al. (2010) conclude that, unless it is combined with communication, monetary punishment is not an effective way to reduce resource extraction. There is also evidence that nonmonetary punishment (Masclet, 2003; Noussair and Tucker, 2005; Dugar, 2010), social approval (Gächter and Fehr, 1999), and public observability (Barr, 2001; Denant-Boemont, 2011; López et al., 2012) can all increase contributions in public good games. Yet it is also demonstrated in Rege and Telle (2004) and in Noussair and Tucker (2007) that initial increases in

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

[2] In line with Bowles and Gintis (2011), by "social preferences" we refer to a wide range of motives such as reciprocity, altruism, and conformism as well as such emotions as shame, guilt, and anger. For the purpose of this study we define "cooperation" more narrowly as the behavior through which one agent internalizes some of the externalities that he imposes on other users, maintaining his own use below levels that would maximize his individual profits.

cooperation induced by public observability tends to fade in the context of a repeated game. We are not aware of any previous evidence concerning the extent to which nonmonetary punishment promotes cooperation in the context of a CPR dilemma.

A number of studies suggest that individuals achieve greater levels of cooperation with members of their own group than with outsiders. Turner et al. (1979) define in-group bias as those instances of favoritism which are unfair or unjustifiable in the sense that they go beyond the objective requirements or evidence of the situation. In this way, individuals enhance their social identity by taking decisions that are more favorable to their in-group than their out-group members (Tajfel and Turner, 1979).[3] This phenomenon has been observed not only in groups induced artificially (Charness et al., 2007; Chen and Xin, 2009; Hargreaves et al., 2009; Harris et al., 2012) but also in groups that occur naturally (Bandiera et al., 2005; Miguel and Gugerty, 2005; Bernhard et al., 2006; Ruffle and Sosis, 2006; Goette et al. 2012). These results are clearly relevant to any analysis of a CPR dilemma. Regimes of property rights granted to one group versus several groups may influence the social preferences of group members and thereby affect resource conservation.

We perform a framed field experiment (Harrison and List, 2004) in which the subject pool consists of fishers on the Uruguayan seacoast who ply their trade in two coastal lagoons while living in nearby villages.[4] Thus our study employs naturally occurring groups in a field setting. We seek to establish whether (or not) fishers who live in different communities are more sensitive to NMP when interacting among themselves than when interacting with fishers from a different community. We also test whether their propensity to cooperate differs in these two scenarios. Fishers from different communities do not interact during their daily life, but they often encounter each other while fishing as they move from one lagoon to the other in pursuit of more available fish. Our experiment incorporates both an NMP and an in-group/mixed-group treatment. Individuals participate in a CPR game and, after five periods of this game, the NMP is implemented. This nonmonetary punishment enables individuals to express their disapproval of others' extraction decisions while facing a monetary cost themselves. Disapproval is registered by flags that vary in color to reflect the level of disapproval. For the in-group treatment, subjects interact only with members of their own community; for the mixed-group treatment, subjects interact also with members of another community.

The paper contributes to the literature in bringing the devices of social punishment from the lab to the field and replicating some of the usual findings found on the former. It combines three features that have not previously been implemented at the same time. First, instead of inducing artificial in-group/mixed-group differences, our setup involves individuals from actual separate communities meeting each other. As pointed out by Cardenas (2003) and Bernhard et al. (2006), naturally occurring groups provide an ideal environment for the study of how group affiliation

[3] More broadly, Akerlof and Kranton (2000; 2005) and Bowles and Gintis (2002) have highlighted the relevance of social identity and group affiliation to the behavior of individuals in most economic organizations.

[4] We concentrate on coastal lagoons because they are exploited only by artisanal fishers—in contrast to open sea, where large-scale fishing is widespread.

affects social norms. Second, groups are reshuffled after each period in order to avoid repeated game effects that could lead to a self-sustaining cooperative equilibrium. Third, even though individuals who are socially punished incur no monetary cost, those who punish others *are* assessed a monetary cost; this protocol was implemented to reduce the likelihood of carelessly administered punishments.

We find that nonmonetary punishment has a positive effect on cooperation when individuals are interacting with fishers from other communities. That is, when individuals in the mixed-group treatment face the possibility of NMP, they reduce their average extraction level irrespective of whether they are actually punished. The effectiveness of these informal sanctions is compromised to the extent that some individuals are less sensitive than others to NMP and also by the use of such sanctions to punish not only free riders but also cooperators. We observe that individuals adjust their extraction levels from one period to the next in order to converge to the previous period's group average—suggesting, in effect, that conformity is the prevailing social norm. Also, subjects whose partners were punished in the previous period then prefer behaving less cooperatively to running the risk of being disadvantaged by others' decisions. This dynamic confirms the notion that social preferences depend on the context. We conclude that the threat of disapproval induces more cooperative behavior only in mixed-group scenarios. This finding suggests that cooperation could be enhanced by implementing social sanctions as a way to govern CPRs that are exploited by more than one community. Yet because social norms need not be shared by all individuals, we conclude that peer punishment must be coordinated in order to prevent antisocial targeting and to enhance the social signal conveyed by punishment.

The rest of the paper is organized as follows. In Section 2 we describe the experimental design. Section 3 reports our results, and we conclude with a summary discussion in Section 4.

2. Experimental design

2.1. Subject pool

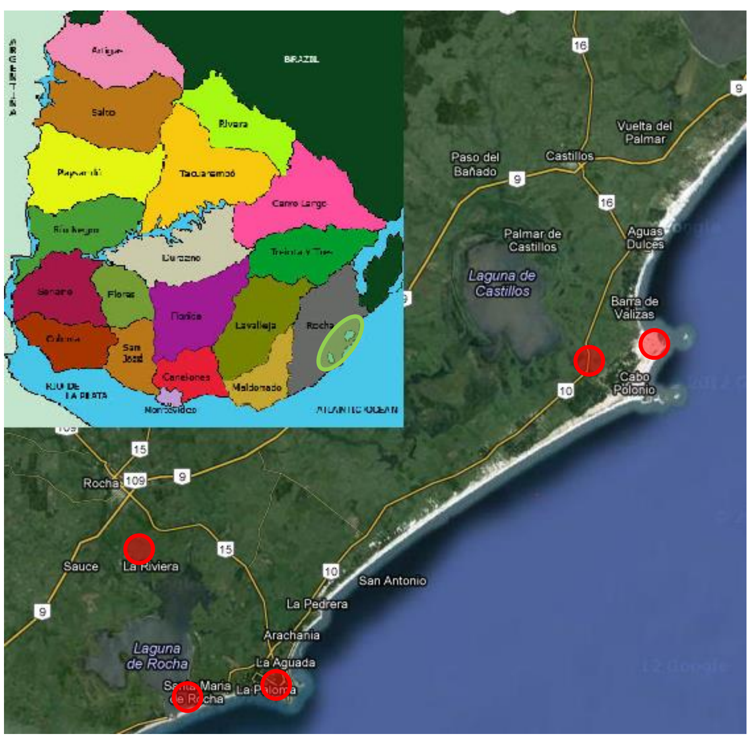
The Laguna de Rocha and Laguna de Castillos are located in the south-east coast of Uruguay (Figure 1). The former was declared part of the national protected areas system in 2010, while the latter is in the process of being declared a protected area. Artisanal fishing activities are mostly developed by fishers from local communities, who move from one fishing site to the other depending on the season (CAEAPLR, 2006; Rodríguez-Gallego et al., 2008; and Defeo et al., 2009). Fishers have to get a specific permit from the National Aquatic Resources Direction (DINARA) in order to be able to fish in the coastal lagoons, and the size and the number of fishnets per license is regulated. The DINARA conducts inspections, and violations of regulations can be punished with sanctions that range from fines to gear confiscation. In practice, licenses work more as an administrative register than as mechanism to regulate access. Also, inspections are ineffective because of institutional weaknesses.

There are no rules that grant privileges to local fishers, any person holding a permit for that specific zone (coastal lagoons, their tributaries and the Atlantic Ocean) is allowed to fish. Fish overexploitation is one of the most visible pressures that the Laguna de Rocha suffers, with some species exhibiting poor reproduction dynamics (Rodríguez-Gallego et al., 2008 and Defeo et al., 2009). This is a major concern, since fish are essential for both the conservation of local ecosystem and the preservation of the main source of income for the mostly low-income residents (Thompson, 2007).

From five communities we recruited individuals who fish in the Laguna de Rocha and/or in the Laguna de Castillos, two coastal lagoons about 50 kilometers apart on the Uruguayan seacoast; see Figure 1. We define community as a group of people who live in the same settlement and constantly interact with each other. Individuals from different communities are ethnically homogeneous but exhibit differences in socioeconomic characteristics. These subject communities differ in terms of how connected they are to relevant markets and also in terms of their exit options. The members of some communities (Laguna de Rocha, Puerto los Botes, and—to a lesser extent—El Puente) lead extremely isolated lives, and fishing is their main source of income; those in the other communities (Valizas and Barrio Parque) are more connected to densely populated areas and so have more varied options. Individuals with more exit options typically have more income and greater wealth (see Table A.1 in the Appendix).

Fishermen from different communities seldom meet in their daily lives, but they do so more frequently when moving across lagoons while fishing in seasonal peaks. Such movement is prevalent during the peak shrimp season, which usually occurs at least once annually in the Laguna de Castillos but rarely (for geographical reasons) in the Laguna de Rocha. According to PROBIDES (2002), fishermen have complained about fishers from other communities who arrive during the peak season to fish in the lagoon where the complainants fish year round. We believe that place of residence is the main factor dividing groups of fishermen.

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



2.2. The experiment

Our experiment consisted of a 20-period CPR game structured in two stages of 10 periods each. In both stages, subjects either interacted only with members of their own community (in-group treatment) or interacted also with fishermen from another community (mixed-group treatment). This distinction was *not* explained to the participants.

The CPR game was used to frame the decision of how many nets to use when fishing. Subjects made their decision in subgroups of four participants. During the first five periods of each stage, subjects participated in a regular CPR game in which they considered a common pool resource exploited by individuals who have the same maximum endowment (eight nets) of fishing rights. An individual's benefits were increasing in the number of nets he used and decreasing in the aggregate number of nets used (see Table A.2). Subject i's earnings in each of the *first* five periods (of each stage) are given by the payoff function $\pi_i = 18a_i + 12\sum_{j=1}^{4}(8 - a_j)$.[5] A selfish individual would always choose $a_i = 8$ so as to maximize his own material payoff. We interpreted any deviation from that choice as an indication of the individual's social preferences.

During the last five periods of each stage, a nonmonetary punishment was introduced. During these periods, subjects were allowed to express disapproval of others' fishnet choices. Those who were punished by others were issued a flag whose color (yellow, orange, or red) indicated the extent (least to most, respectively) of their peers' disapproval. After deciding how many nets to use and then being informed of the total number of nets used by other subgroup members (which enabled calculation of an average extraction value), each participant was empowered to allocate 0 to 10 "punishment points" to the number of fishnets that others chose to use (see Table A.3). A subject could disapprove of eight extraction alternatives (1 to 8 nets) at the same time.[6]

Punishment points resulted in no monetary cost to the punished but did entail a monetary cost to the punisher: each punishment point cost the punisher one point in his "earnings account". Subjects were charged for the total number of punishment points they issued regardless of whether that number corresponded to the number of nets actually used by the targets.[7] Thus

[5] Unlike the quadratic model representing decreasing marginal returns common in CPR framed field experiments (e.g. Ostrom et al., 1994; Cárdenas, 2003), we specify a linear payoff function. The model in this article is an adaptation of the model developed by Bowles (2004) in which peer monitoring and forms of social disapproval enable individuals to achieve agreed levels of effort. We choose the payoff function to be linear in order to simplify the model given the greater complexity of considering motives for social disapproval. The theoretical model guiding this article is available in a working paper version (http://ideas.repec.org/p/ulr/wpaper/dt-16-12.html). The aim of the payoff function is to test for the existence of cooperation.

[6] A subject was free to punish those choosing the same number of nets that he chose, although such punishment would not also be applied to himself (this aspect of the setup was explained only in response to a direct question).

[7] There are two reasons why the punisher was charged the total number of disapproval points and not merely for those corresponding to actual fishnet choices: the former (i) was much simpler to explain and (ii) enabled the subject to calculate the cost by himself. We believe that simplifying mechanisms of this type are especially useful in a framed field environment.

subject *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 much lower than the points a subject could *earn* during a period. For instance, if all subjects chose the Nash equilibrium in one period then each would earn 144 points; if the social optimum was achieved, then each would earn 354 points. The cost to a subject who disapproved of all possible fishnet choices by administering the maximum punishment in one period would total 80 points (equivalent to about half a US dollar, hereafter denoted simply via "$"). The aim of this treatment was to re-create the experience of being socially punished in the field (via gossip, direct criticism, etc.) and to evaluate the effects of that punishment on extractive decisions in subsequent periods. We acknowledge that punishing others socially may entail also a social cost to the punisher, but because that is not the focus of this study it is sufficient for our purposes that there be *some* (monetary) cost to the punisher. Carpenter (2007) concludes that the demand for punishment is relatively inelastic with respect to price and income and argues that this is due to the fact that individuals punish primarily for social rather than economic reasons.

The punishment points administered to each subject were totaled, and yellow, orange, and red flags were then assigned in accordance with the ranges listed in Table A.4. No subject could receive a red flag unless more than one other participant disapproved of his fishnet choice.

We employed a *hybrid* strategy method to implement this treatment. 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. In order to avoid these shortcomings, we chose a *hybrid* strategy method, one that is closer in spirit to assigning punishment based on knowing the fishnet choices of each of the other subgroup members, but that still preserves anonymity. In this *hybrid* strategy method individuals made decisions in two stages (and not as in the classical procedure whereby both decisions—namely, extraction and punishment—are made at the same time). Punishment points were assigned after subjects were informed of the total number of nets used by other subgroup members (allowing each subject to calculate the average number of nets used).[8] 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]

2.3. Structure of the experiment

Subjects were first contacted during a survey conducted in March 2011. The aim of this survey was to gather data on socioeconomic characteristics and environmental perceptions among the resource users of artisanal fisher communities in Rocha and Castillos coastal lagoons. At the end

[8] It is not a pure strategy method because the player already has information about actual extraction choices in the current round when he/she is given the possibility to punish for every possible number of fishnets.

[9] A priority of our research was to maintain anonymity with respect to individual extraction and punishment decisions. Indeed, Anderies et al. (2011) underscore the importance of preserving such anonymity when working with communities in field experiments: because the game may not end when experimenters leave, untoward disclosure may well have negative spillover effects on participants' subsequent daily lives.

of the questionnaire, each interviewee was asked whether he would like to participate in an activity where he could earn, on average, the equivalent of two days' wages (about $30). The experiment´s recruitment took place a week before the experiment. We revisited the five communities and hand-delivered flyers to residents, and we also made phone calls to those who had been surveyed but could not be located during our subsequent visit.[10]

The experiment was conducted in two sessions during November 2011. Both sessions took place at La Paloma, a town in the province of Rocha, Uruguay. The communities that participated in each session were determined randomly (Table 1). Unlike most framed field experiments, in this study subjects were transported from their place of residence to the town in which the experiment was conducted.[11] This design was necessary so that subjects from different communities could meet, but it required that fishermen leave their community to participate. We had difficulty convincing subjects to travel, which explains why there were fewer participants than desired.[12]

When subjects arrived at the venue, they drew a number from a bag (one bag per community). This number represented an identifier that assigned each subject to a group of either eight or twelve members for each stage. Within these groups, participants took part in a CPR game in subgroups of four. In the mixed-group treatment, each subgroup consisted of two individuals from each of two communities.[13] In order to minimize repeated game type of behavior, after each period the subjects were reshuffled into new groupings of eight or twelve. The subgroups that participated in the experiment's 20 periods were predetermined by identifier numbers. It was common knowledge that the matching procedure for all periods was random and was determined by the initial draw of participants' identifier numbers. After each period, the experimenters indicated to the participants which subgroup of four they would take part in the next period. After the first 10 periods, participants in the mixed-group treatment were switched to the in-group treatment (and vice versa), although the subjects were not informed of these particulars. During session 1, the in-group treatment preceded the mixed-group treatment; in session 2 the order was reversed (see Table 1). This design enabled us to control for order effects.

[10] To test for the existence of a self-selection bias between those individuals that participated in the experiment and those that were surveyed in March but did not participate in the experiment, we conducted Wilkoxon rank-sum tests considering the variables in Table A.1. Difference in the means was rejected for all the tested variables (Years of schooling, z= -1.48 p-value= 0.14; Wealth, z=0.059, p-value=0.95; Per capita income (US), 0.471; p-value=0.95; Fishing as the main activity, z=-0.34, p-value=0.74).

[11] Buses were hired to transport participants to and from the experimental venue.

[12] We were unable to expand the sample size by extending the experiment to other communities because the mixed-group fishing scenario that takes place in the Rocha and Castillos' coastal lagoons is almost unique in Uruguay.

[13] In session 2, there were twelve subjects from one community (El Puente) and only eight subjects from the other two communities (Puerto los Botes and Barrio Parque). Hence the mixed-group treatment in this session involved subgroups composed either of two subjects from El Puente and two from one of the other communities or of three subjects from El Puente and one from another community. In all cases, the mixed-group treatment involved mixing just two communities.

Table 1: Characteristics of the experimental sessions

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

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

[b] 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.

Once divided into subgroups of four members, participants were asked to sit back-to-back so that they could not see the others' choices. Each session was conducted by a moderator who gave instructions throughout the game, and each also included a monitor for every subgroup of four. This protocol ensured that subjects did not interact during the game and that an experimenter was always available to explain how the materials should be used.

Subjects received a payoff table and an earnings sheet on which they kept a record of their decisions and points gained. The table summarized the payoff consequences of all combinations of subject's own nets used and the total number of nets used by the subgroup's other three members (see Table A.2). The exchange rate was set at $0.62 for 100 points. Subjects were asked to decide—while looking at the payoff table—how many nets to use (minimum one, maximum eight); this number they wrote on a slip of paper that was then handed to the experimenter. Once the four subjects had transcribed their decisions, the total number of nets used by the subgroup was announced so that each subject could calculate the number of points earned and write that figure on his earnings sheet. The explanation of the game followed that described in Cardenas (2003). The actual experiment began after the moderator had conducted three rehearsal periods and once all questions from participants had been answered. All decisions were made privately and individually, and only the total number of nets chosen (by the four subgroup members) was announced publicly.

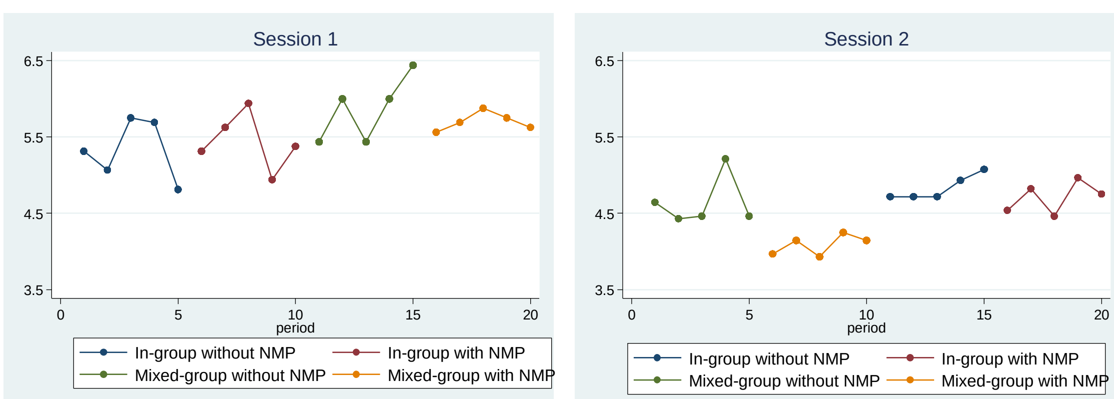
Prior to beginning the experiment's punishment phase (i.e., before the last five periods of each stage), an illustrative example was described to the participants.[14] This example showed three subjects' disapproval cards: one punishing without any criteria, one punishing those who used many nets, and one not punishing at all. Upon resumption of the actual experiment, the subjects' chosen number of nets and number of assigned punishment points assigned were private information; the only public information was the flag received by subjects who received punishments from others of more than one point. Subjects had to display any flag received so that others could see it during the game's next period.

At the end of each experimental session, we conducted a post-experiment survey containing questions about reasons for punishment and about feelings in response to being punished. Each session of the experiment lasted about three hours, and participants earned, on average, nearly $30 (including a guaranteed $5 participation fee); this amount is equivalent to 10% of a subject's typical monthly wage.[15]

3. Results

The left and right panels of Figure 2 plot average extraction levels (number of nets chosen) by period and treatment for sessions 1 and 2, respectively. At first glance, the figure suggests that the change from in-group to mixed-group treatment does not lead to significant behavioral changes in the absence of nonmonetary punishment. Without NMP, the extraction levels chosen by subjects in session 1 were higher for the mixed-group than for the in-group treatment; however, no substantial change was observed for the subjects in session 2. The inclusion of NMP had a positive effect on cooperation, especially for the mixed-group treatment: it lowered average extraction levels in the second stage of session 1 and in both stages in session 2. It is noteworthy that, in session 2, members of the three participating communities chose extraction levels that were significantly lower than those chosen by members of the two communities participating in session 1.

Figure 2: Average extraction levels by treatment type and session



[14] Subjects were neither told that in rounds 6-10 and 16-20 they would be allowed to disapprove of others´ behavior nor that in rounds 11-20 they would change groups completely (groups of 8-12). They were told that they would switch groups of 4 every single round.

[15] A similar experimental design (which excluded the in-group and mixed-group treatments) was tested using 36 undergraduate students as participants.

3.1. Testing the treatment effects

To test the effects of our in-group/mixed-group and NMP treatments, we employed dynamic analysis to examine the extractive decisions of participating subjects. These treatments were tested in two ways. First, the model included a pair of indicator variables: *in-group*, which was set equal to 1 (resp., to 0) for subjects when part of the in-group (resp., mixed-group) treatment; and *NMP*, which was set equal to 1 for extraction decisions made during a round that allowed for NMP (i.e., rounds 6–10 and 16–20) and to 0 otherwise. Second, we tested for interaction between the treatments. For this purpose, three dummy variables were included: *mixed-group with NMP*, *in-group with NMP*, and *in-group without NMP* (recall that the base-case scenario is "mixed-group without NMP"); each of these dummies was set equal to 1 only for the scenario that it describes, and 0 otherwise.

A fixed-effects regression was performed that controls for individuals' time-invariant characteristics. The final model estimated (model [8] in Table 2) is

(1) $$a_i^t = \alpha_i + \beta_1 mixed\_group\ w/NMP_i^t + \beta_2 in\_group\ w/\ NMP\ _i^t + \beta_3 in\_group\ w/out\ NMP\ _i^t + \beta_4 \pi_i^{t-1} + \beta_2 \pi_{-1}^{t-1} + \beta_5 stage_i^t + eit$$

In this formula, a_{it} denotes *i*'s extraction level in period *t*. Equation (1) also includes, as additional controls, two variables that are frequently used in the literature: $\pi_{i;\,t-1}\boldsymbol{\pi_{i;t-1}}$, which is the payoff to individual *i* from the previous round; and $\pi_{-i;\,t-1}\boldsymbol{\pi_{-i;t-1}}$, the payoff from that round to the rest of individual *i*'s subgroup (i.e., excluding *i*). High payoffs occur as a result of either cooperation (high group payoff and high individual payoff) or self-interested behavior (low group payoff but high individual payoff). Controlling for the group's payoff allows us to distinguish which of the two strategies is reinforced over time. Even though the game is a series of one-shot rounds and the subgroup members change every period, a subject can use information about the behavior of other subjects to guide his own future behavior. A negative relation between the group's previous-period payoff's and an individual's extraction levels suggests the existence of social preferences. Finally, we include an indicator variable, *second stage*, that is set equal to 1 only for rounds 11–20. The last term in equation (1) is a normally distributed random residual. Note that time fixed effects were omitted because they are strongly correlated with the treatment variables; in fact, our treatment dummies are themselves time.

Columns [1]–[6] of Table 2 show that, whereas the in-group treatment has no effect on subjects' choice of extraction level, lower levels are chosen under the NMP treatment. The values reported in column [2] confirm that the NMP treatment effect is significant irrespective of which additional variables are included.

Columns [7] and [8] of this table document that the treatment effects of NMP differ for in-group versus mixed-group settings. On the one hand, the number of nets chosen in the mixed-group setting without NMP is not significantly different from the number chosen in the in-group setting—with or without nonmonetary punishment. On the other hand, subjects in the mixed-group setting choose fewer nets with than without NMP (the −0.4 coefficient amounts to 20% of a standard deviation in the number of nets), yet the behavior of individuals in the in-group setting is not significantly affected by NMP. Finally, the *second stage* dummy variable is positive and significant in all models. In other words, subjects increased their average extraction level during the second stage regardless of which treatment they experienced first. This finding—that cooperation decays throughout the game—is in line with previous research.

Table 2: Dynamic analysis of extraction 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)		
mixed--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)
Obs.	880	880	880	836	836	836	880	836
Subjects	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

*** p<0.01;** p<0.05; * p<0.1

Robust standard errors in parenthesis

As regards the earnings of participants in previous rounds, models [4], [5], [6], and [8] exhibit a positive coefficient on individual i´s payoff and a negative coefficient on the payoff of the rest of individual *i*'s subgroup. This result is consistent with Hayo and Vollan (2012) and suggests that individual decisions are influenced by social preference mechanisms.[16] The negative coefficient on the payoff of the rest of individual *i*'s subgroup implies that individuals behave

[16] We include individual *i's* payoff in a previous round and the payoff of the rest of individual *i*'s subgroup (excluding individual *i*) in the previous round as controls. Following Hayo and Vollan (2012) 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. 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 reciprocity.

more cooperatively if their group performed well in the previous round. That is, an individual's recent past experience affects his choice of extraction level even though his partners change after each round.

3.2. Determinants of extraction decisions

In this section we address the question of whether there are sociodemographic determinants of individual extraction decisions. The analysis is performed for three variables of interest: number of nets chosen in the first period (columns [1] and [2] of Table 3), total number of nets chosen throughout the 20 periods (columns [3]–[5]), and average number of nets (columns [6]–[8]). The table reports the general and reduced estimates for each of these variables.

Wealth and age are the only observable individual determinants that are statistically significant; no other individual-level economic or demographic variable is able to explain extraction choices. These findings echo those of Henrich (2001) and Hayo and Vollan (2012). The magnitude of the wealth coefficient is worth noting: an increase of one standard deviation in the wealth index increases by 44% the average number of nets chosen (the wealth index, which is based on the durable goods owned by a household, was elaborated by means of factor analysis).[17] Cardenas (2003) also finds a positive relation between wealth and choices of extraction levels, and he hypothesizes that low-wealth status may reflect greater experience in managing a common pool resource. Yet this does not seem to be the case in our study, as being a fisherman as one's main activity bears no relation to extraction levels (see Table 3). Cardenas also provides an alternative explanation that could apply to our case as follows. Wealthier participants may gain less marginal utility from the cash earned in the experiment; that is, they have less incentive to cooperate because the marginal value of potential gains is smaller than for poorer participants. Hayo and Vollan (2012) report a positive coefficient for the two highest quartiles of income, arguing that high income may translate into stronger preferences for consumption, risk, and competition.

One 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 (WMW) ranksum test rejects the equality of median and mean extraction levels between places of residence, two-by-two, at the 10% confidence level.[18] These results—together with the non-significance of individual characteristics—strongly support the hypothesis that group-level institutions and/or social norms affect the behavior of individuals.

[17] This index includes variables for the following goods: water heater, refrigerator, television, radio, cable TV, DVD player, washing machine, microwave, computer, Internet connection, telephone, motorbike, automobile, and horse.

[18] There are only two cases in which this hypothesis is *not* rejected: (i) when comparing Barra de Valizas and Barrio Parque with respect to the average number of nets used; and (ii) when comparing Laguna de Rocha, Barra de Valizas, and Barrio Parque with respect to the average number of points earned during the experiment. However, equality between Barra de Valizas and Barrio Parque with respect to median average earnings is rejected by the WMW ranksum test. Tests are available upon request.

Table 3: Determinants of subjects' extraction decisions

	Dependent variable							
	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)	(13.96)	(9.23)	(9.76)	(0.70)	(0.46)	(0.49)
Valizas	2.53**	1.17	52.92***	51.88***	50.16***	2.65***	2.59***	2.51***
	(0.93)	(0.85)	(15.33)	(11.80)	(12.92)	(0.77)	(0.59)	(0.65)
Botes	2.97**	1.42*	25.05	24.22**	25.25**	1.25	1.21**	1.26**
	(1.09)	(0.76)	(15.11)	(9.97)	(9.77)	(0.76)	(0.50)	(0.49)
Barrio Parque	3.23**	1.42	38.35**	29.81**	31.23**	1.92**	1.49**	1.56**
	(131)	(0.94)	(18.67)	(13.45)	(13.79)	(0.93)	(0.67)	(0.69)
female	0.64		-5.46			-0.27		
	(0.90)		(7.69)			(0.38)		
age	-0.02		-1.07***	-0.53**	-0.58**	-0.05***	-0.03**	-0.03**
	(0.03)		(0.30)	(0.24)	(0.26)	(0.02)	(0.01)	(0.01)
years of schooling	-0.02		-3.50*			-0.17*		
	(0.17)		(1.80)			(0.09)		
drinkable water	-1.80*		-6.39			-0.32		
	(1.01)		(14.36)			(0.72)		
electricity	-1.04		-17.02			-0.85		
	(1.03)		(14.13)			(0.71)		
wealth[a]	0.49**		11.94***	8.04***		0.60***	0.40***	
	(0.23)		(2.74)	(2.10)		(0.14)	(0.10)	
per capita income (logs)	-0.99**		1.50			0.07		
	(0.43)		(5.91)			(0.30)		
fishing main activity	1.11*		-2.29			-0.11		
	(0.65)		(7.89)			(0.39)		
perception[b]	-0.28		-3.85			-0.19		
	(0.72)		(8.50)			(0.43)		
trust[c]	-0.16		-13.11			-0.66		
	(1.20)		(11.26)			(0.56)		
second quartile (wealth)					7.00			0.35
					(11.85)			(0.59)
third quartile (wealth)					25.35**			1.27**
					(11.27)			(0.56)
fourth quartile (wealth)					27.02**			1.35**
					(11.78)			(0.59)
Constant	11.85***	3.83***	114.58*	72.59***	81.56***	5.73*	3.63***	4.08***
	(3.61)	(0.55)	(58.61)	(11.53)	(13.89)	(2.93)	(0.58)	(0.69)
Obs.	43	44	43	44	44	43	44	44
R-squared	0.34	0.12	0.61	0.46	0.45	0.61	0.46	0.45

Robust standard errors in parentheses

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

[a] The wealth index considers different durable goods a household may own.

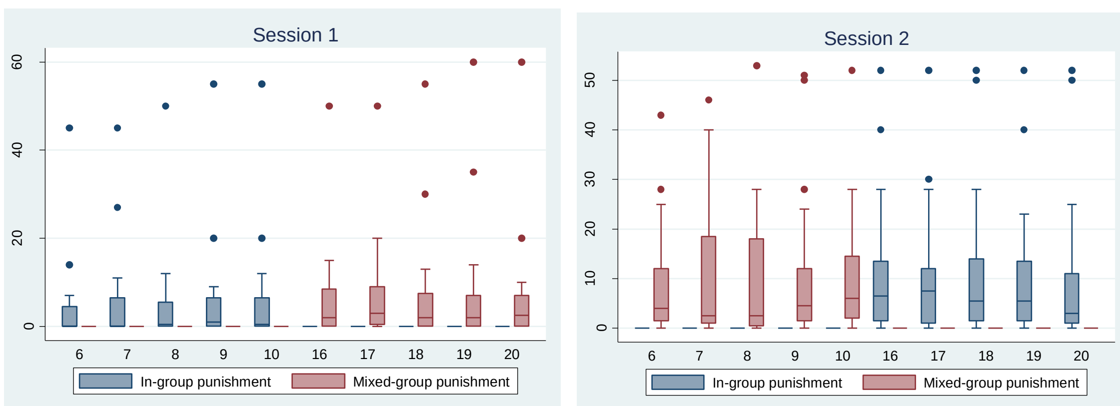
[b] Believes that preserving the environment in coastal lagoons is mainly a responsibility of the people rather than the government.

[c] Believes one can trust most people.

3.3. Punishment behavior

In this section we analyze the behavior of punishers. On average, 71% of the subjects chose to punish during each period in which punishment was allowed. Disapproval was substantial throughout the game and surprisingly high in the last period, even though subjects knew the experiment would be over after that period. Figure 3 graphs the distribution of punishment points administered, by period, for the two sessions. Because subjects were unaware of the individual extraction level of the other subgroup members, during the mixed group treatment they could not knowingly direct punishment to members of the other community. Higher levels of punishment were observed under the mixed-group treatment in session 1 but not in session 2. When both sessions are considered together, there is no statistically significant difference in the amount of punishment administered under mixed-group versus in-group treatments.[19]

Figure 3: Distribution of punishment points by period and session



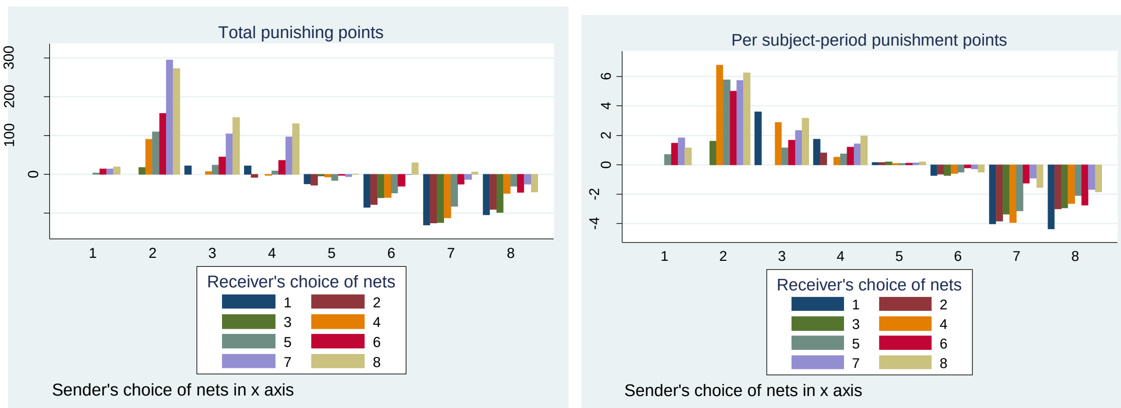
Note: The figure plots the distribution of punishment points for each period and session. The line inside the box represents the median punishment points while the lower and upper hinges of the box represent the 25th and 75th percentiles.

Following Herrmann et al. (2008), we interpret punishment for extraction levels greater than (resp., less than or equal to) one's own as punishment of free riding (resp., as antisocial punishment). Figure 4 displays positive values for punishment points directed at free riders and negative values for antisocial punishment. The left panel of this figure reports total disapproval points in terms of how many nets the punisher (sender) chose (horizontal axis). Bar colors indicate how many nets the punished (receiver) chose. In general terms, subjects who choose fewer than five nets (horizontal axis)assign more punishing points to those who choose six or more nets (red, purple and beige bars). Such punishment of free riding could be viewed as “altruistic” punishment: individuals incur material costs when punishing but reap no material benefits from doing so, since participants are reshuffled after each period. We also observe antisocial punishment (i.e., administered to cooperators), as when those choosing five or more nets (horizontal axis) mainly disapprove of those choosing fewer than six (blue, brown, green and orange bars). The right panel of Figure 4, which displays per-subject rather than total punishment points, shows that the observed antisocial punishment is the consequence of a few subjects

[19] The WMW ranksum test does not reject equality of punishment administered during mixed-group and in-group treatments for the two sessions together ($p = 0.54$).

administering large amounts of punishment: there are only three subjects that used six or more nets, and they administered a large number of punishment points to those who used fewer nets.[20] Such "misdirected" punishment is also observed by Falk et al. (2000), Masclet et al. (2003), and Gätcher and Herrmann (2011). Nonmonetary punishment is actually more effective than indicated by Figure 2 when the subgroups in which these three subjects participated are excluded. Figure 4 also shows that some subjects punish another who uses the *same* number of nets—especially when that number is large. This could be interpreted as an attempt to discourage others from free riding even as the punisher disregards the social norm (i.e., "do as I say, not as I do").

Figure 4: Punishment behavior by receiver's (bars) and sender's (horizontal axis) extraction choices



Note: This figure reports total disapproval points in the left panel (per subject-period disapproval points in the right panel) in terms of how many nets the punisher chose in the horizontal axis; bar colors indicate how many nets the punished chose.

Translating punishment points into flags reveals the actual intensity of NMP received. On average, 1.7 flags were displayed each period per subgroup (of four members). Table 4 reports the distribution of flags as a function of whether the focal subject chose an extraction level below or above his subgroup's mean. Most flags were applied to subjects who deployed more nets than the mean number in their subgroup, but a full 40% of the flags were applied to individuals whose extraction levels were actually below the subgroup mean.

Table 4: Total flags by round

Period	Total flags	Negative deviation $\max\{0;\bar{a}_{t-1}-a_{i,t-1}\}$				Positive deviation $\max\{0;a_{i,t-1}-\bar{a}_{t-1}\}$			
		Total flags	Yellow	Orange	Red	Total flags	Yellow	Orange	Red
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%

[20] As with punishment overall, antisocial punishment does not differ significantly between the in-group and mixed-group treatments.

Among the post-experiment questions was one asking subjects to identify their reasons for punishment. Most respondents (55%) disapproved of subjects for using too many nets.[21]

Next, we analyze the determinants of punishing each extraction level. For this we use the following model, introduced by Masclet et al. (2003):

$$(2) \quad P_{ik}^t = \beta_0 + \beta_1\left(\max\{0, a_i^t - a_k^t\}\right) + \beta_2\left(\max\{0, a_k^t - a_i^t\}\right) + \beta_3\left(\left(\max\{0, a_k^t - \bar{a}_{subgroup}^t\}\right)\right) + \beta_4\left(\max\{0, \bar{a}_{subgroup}^t - a_k^t\}\right) + e_i^t$$

Here P_{ik}^t is the number of disapproval points that *i* assigns to *k* in round *t*. The coefficient β_1 applies to positive deviations from the punisher's fishnet choice (i.e., cases in which the punished chose fewer nets than the punisher), and β_2 applies to negative deviations from the punisher's fishnet choice (cases in which the punished chose more nets than the punisher); the coefficient β_3 (β_4) applies to the effect of positive (negative) deviations from the subgroup's average. As before, the last term is a normally distributed random residual. In this regression we included individual fixed effects to control for subjects' time-invariant characteristics. We estimated equation (2) for each fishnet choice that could be punished. For instance, the first column in Table 5 reports the determinants of punishing subjects who chose one fishnet. The table indicates that both positive (antisocial punishment) and negative (punishment of free riding) deviations from the punisher's fishnet choice are significant. As Masclet et al. (2003) show related to deviations of the punished subject from the subgroup's average,[22]

[21] The other reasons given for punishment were "without any criteria" (14%), "did not disapprove" (11%), "those who threw few nets" (7%), "those who play differently" (5%), "because it was part of the game" (5%), and "did not understand" (5%).

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

Table 5: Determinants of punishment points assessed to each fishnet option

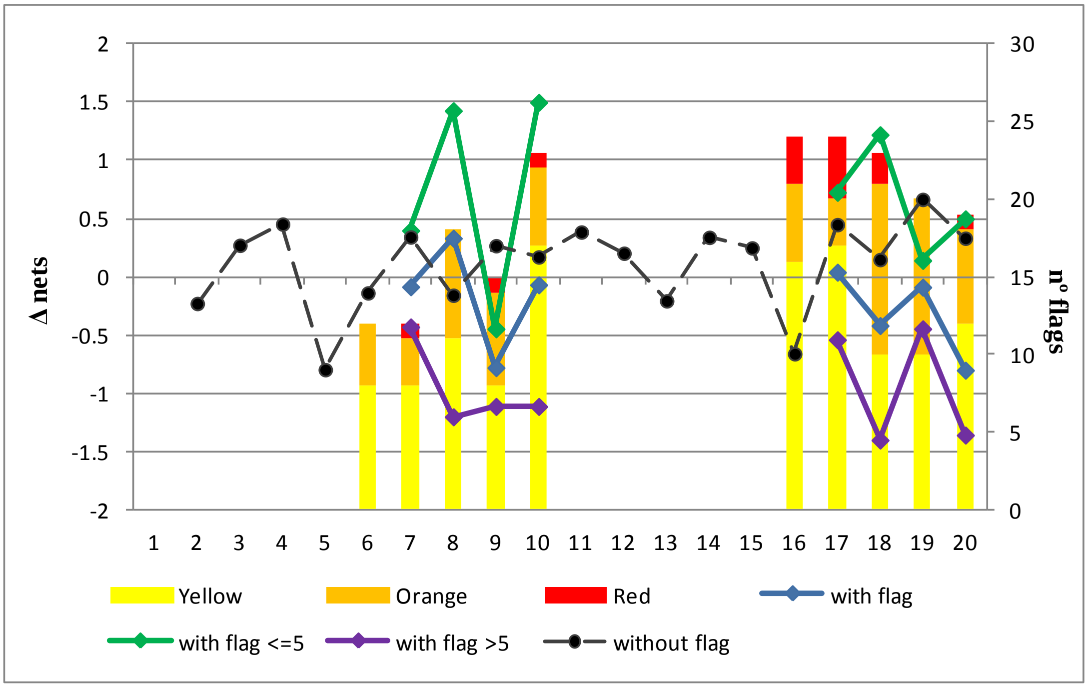
	Dependent variable: Disapproval points							
	Others´ fishnet options							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Positive deviation from i's own extraction ($max\{0, nets_i-nets_k\}$)	0.79***	0.53***	0.89**	0.48**	0.67***	2.22***	1.46**	
	(0.25)	(0.16)	(0.36)	(0.20)	(0.22)	(0.73)	(0.69)	
Negative deviation from i's own extraction ($max\{0, nets_k-nets_i\}$)				1.33**	0.35***	0.49**	0.33**	0.04
				(0.57)	(0.08)	(0.18)	(0.13)	(0.09)
Positive deviation from average ($max\{0, nets_{av}-nets_k\}$)	0.13	0.31	0.03	0.83***	0.70*	0.54	0.31	
	(0.24)	(0.27)	(0.57)	(0.25)	(0.37)	(0.43)	(2.03)	
Negative deviation from average ($max\{0, nets_k-nets_{av}\}$)			-0.43	-0.16**	0.32*	0.26**	0.29**	0.62***
			(0.46)	(0.08)	(0.18)	(0.11)	(0.14)	(0.13)
Constant	0.24*	0.28***	0.30***	0.42***	0.49***	0.38***	0.84***	0.98***
	(0.14)	(0.08)	(0.10)	(0.07)	(0.05)	(0.11)	(0.19)	(0.21)
Obs.	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

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.4. Reaction to punishment

In this section we analyze whether punishment generates a change in behavior among those who are punished. At first glance, the descriptive analysis suggests that flags induce variations in the behavior of individuals. As shown in Figure 5, subjects who receive a flag in one period reduce their extraction in the next period by 0.26 nets on average. This result does not always obtain at the individual level, however, and one reason is that sometimes even those who choose a *low* number of nets are punished. Figure 5 shows that, whereas those who receive a flag after throwing more than five nets reduce that number (by 0.99 nets on average) in the next period, those who receive a flag after throwing five or fewer nets actually increase that number (by 0.61 nets on average) in the next period. In the periods during which NMP was implemented, the number of nets chosen by those who do not receive a flag in the previous period ranges around zero.

Figure 5: Fishnet variations and total number of flags



A significant percentage (33%) of the individuals who received a flag did not change their behavior in the next round. The modes of the distribution of extraction choices were two and eight nets, and nearly half (55% and 48%, respectively) of subjects who chose those values did not change their choice—that is, irrespective of how many flags they received. Thus cooperation may not be the norm that punishers aim to enforce; there may be other norms (e.g., "try to fish as much as possible") that may prevail (Noussair et al., 2011). Hence some punished subjects may view as inappropriate such disapproval for using many nets and respond by increasing their number of nets (or maintaining their choice of the maximum number).

Not all the flag colors yielded the same reaction. Subjects were more indifferent to yellow than to other flags: in 42% of the cases where a subject received a yellow flag, he did not change his decision in the next period. When analyzing subjects' reactions in view of their feelings (as described in the post-experiment survey) after receiving a flag, we find that—of those subjects who declared indifference to a flag—70% either maintained or increased their extraction level after receiving one. In contrast, those subjects who admitted to feeling uncomfortable when flagged either maintained (in 28% of the cases) or reduced (in 52%) their extraction level during the next period after being flagged. All subjects who responded that they experienced anger when punished increased their extraction level in the following period.[23]

The next step is to devise a formal test for the behavior just described. We first test for whether subject i's decision changes from period $t-1$ to period t as a function of punishment received in

[23] To save space, we have omitted tables reporting extraction level variations as a function of flag color received and of subjects' self-reported feelings in response to punishment. These tables are available from the authors upon request.

the previous period. Then we adapt the reaction function employed by Masclet et al. (2003) and Noussair and Tucker (2005) and test for whether—once the subject's deviation (from his subgroup average) is included in the regression—any such decision changes are still related to punishment received in the previous period. The model estimated is:

(3) $$a_i^t - a_i^{t-1} = \beta_0 + \beta_1 * Flag_i^{t-1} + \beta_2 * OthersFlag_i^{t-1} + \beta_3 * \left(\max\left\{0, a_i^{t-1} - \bar{a}^{t-1}\right\}\right) + \beta_4 * \left(\max\left\{0, \bar{a}^{t-1} - a_i^{t-1}\right\}\right) + e_i^t$$

Here $Flag_i^{t-1}$ is a dummy variable that indicates whether or not the individual received a flag in a previous period, and $OthersFlag_i^{t-1}$ denotes how many of i's partners in period *t* received a flag in period $t-1$ (this variable ranges in value from 0 to 3). The terms $max\left\{0, a_i^{t-1} - \bar{a}^{t-1}\right\}$ and $max\left\{0, \bar{a}^{t-1} - a_i^{t-1}\right\}$ capture whether the subject extracted (respectively) more or less than his subgroup's average in the previous period and also indicate the magnitude of any deviation. Once again, the equation's last term is a normally distributed random residual.[24] Regression results for this model are reported in column [3] of Table 6. We test this model only for those periods during which reactions to flagging could occur (i.e., periods 7–10 and 17–20) and separately for subjects who chose five or fewer nets and for subjects who chose more than five nets. The table also includes (as column [11]) results for a control model that applies to periods during which flag reactions were not possible; this allows us to compare conformity effects. All specifications include individual fixed effects to control for non-observable factors that may affect a subject's decisions.

[24] We also estimated alternative specifications in which flag colors were distinguished; the results from these estimations are substantially similar to those reported here.

Table 6: Reaction to punishment

	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-1	-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

The values reported in column [1] of Table 6 indicate that being punished results in a downward adjustment during subsequent periods. Columns [2]–[5] and [7]–[10] present results when the model includes a count variable for how many of the three other members in a subject's group during period t received a flag during period $t-1$. Having partners who were punished in the previous period tends to increase own extraction, an effect that is more pronounced for subjects who chose to deploy five or fewer nets in the previous period (columns [4] and [9]).[25] However, such subgroup partners have no effect on an individual who used more than five nets in the previous period (columns [5] and [10]). These findings lead us to conclude that social preferences are context dependent: a subject who sees himself among noncooperative partners reacts by increasing his own extraction level to avoid being disadvantaged. In this sense, individuals behave as if punished subjects will not react favorably (i.e., by reducing their level of extraction) to that punishment. Participants may also interpret that they are less likely to be punished for using many nets when their subgroup partners are accustomed to extracting high levels of the resource.

If we adjust for conformity effects—that is, for the deviation of individuals from the average of their subgroup in the previous period—then receiving a flag is no longer followed by reduced extraction levels (columns [3]–[5] and [8]–[10] of Table 6). Those who used fewer (more) nets than the average of their previous period's subgroup will increase (decrease) their extraction in the next period. These conformity effects are consistent with the results reported by Masclet et al. (2003), Velez et al. (2009), and by Hayo and Vollan (2012). As might be expected, the number of nets does not decrease when we consider only the reaction of those who received a flag after choosing five or fewer nets and, conversely, the number of nets does not increase when we consider only the reaction of those who received a flag after choosing more than five nets. The magnitude of the conformity effect is greater during the NMP periods (especially for positive deviations in subjects' extraction levels relative to their subgroup's mean), which could suggest that nonmonetary punishment is inducing convergence to the social norm. However, the confidence intervals for the effects described here, in periods with and without NMP, overlap at the 95% level (columns [11] and [12]). Interactions between previous-period deviating behavior and being flagged are not significant, which confirms that the high significance of conformity effects is not a consequence of being punished. Note also that these conformity effects do not differ for in-group versus mixed-group treatments (see Table A.5).

When flags are distinguished by their colors (column [6] in Table 6), we observe that receiving an orange flag has a modest reducing effect on extraction levels, but this effect is diminished when we split the sample between subjects who were flagged after choosing five or fewer versus more than five nets (columns [9] and [10]). Note the large increase in number of fishnets deployed during period t when a red flag was received after deploying five or fewer nets during period $t-1$ (column [9]). In other words, subjects react strongly—and non-cooperatively—in response to what they perceive as unfair punishment.

In short, nonmonetary punishment (as explained in Section 3.1) reduces extraction levels—especially for the mixed-group treatment. Yet the period-by-period variations in number of nets used reveal that individuals adjust their choice of extraction level mainly in response to their subgroup's average in the previous period and *not* in response to punishment. This conformity

[25] Recall that punishment entails displaying any flag received in one period during the game's next period.

effect is evident regardless of the availability of NMP. The only participants who respond to received punishment by choosing to deploy *more* nets are those who receive a red flag and believe that this punishment is unfair. That being said, subjects increase their own extraction if their current partners were punished in the previous period. Participants do not expect a punished subject to respond by reducing his level of extraction—to the contrary, any flag received is interpreted as a signal that he will extract high levels in the next period. Overall, then, subjects adapt their behavior not only to the norms exhibited by their previous round's subgroup but also to the signals of members in their current subgroup.

4. Discussion

In this study we performed a framed field experiment to test the effectiveness of nonmonetary punishment (NMP) in the context of a common pool resource game. We combined this treatment with an in-group/mixed-group treatment requiring fishers from different communities to interact solely with members of their own community in one of the stages of the experiment and mixed with subjects from another community in the other stage.

First, our findings suggest that NMP has the effect of diminishing extraction levels, but only in the mixed-group treatment. That is, at least when interacting with subjects who are not of their own community, subjects achieve greater cooperation levels upon the threat of being punished. In contexts where individuals do not know each other (or hardly know each other) but are aware that there is some chance of seeing each other again, public punishment might constitute the only information others have about oneself so in this sense it is important to avoid being "flagged" via NMP. Such punishment may not be perceived as meaningful when administered by workmates or neighbors, and neither would it matter much when individuals are certain that they will never meet again. In short, the relationship between sensitivity to peer punishment and social context—that is, in-group versus mixed-group interactions—may not be a monotonic one. Previous literature addressing contributions in public good games has found that nonmonetary punishment increases cooperation (López et al., 2012), but it has less of an effect than monetary sanctions (Masclet et al., 2003) and is more effective in increasing cooperation when *combined* with monetary sanctions (Noussair and Tucker, 2005). This paper is consistent with those studies in finding that nonmonetary punishment can enhance cooperation simply by affecting prosocial emotions, yet this result holds only when the subjects belong to different groups.

Second, once we control for conformity effects, reactions to punishment (i.e., reducing extraction levels after receiving a flag) are no longer significant. One reason for this failure of punishment may be that cooperation (here, choosing to use fewer nets) may not be perceived as a social norm enforceable by punishment (Casari and Luini, 2009); thus, there may be other norms—for example, "try to catch as many fish as possible"—that may prevail (Noussair et al., 2011). Also, there are aspects of the subjects' daily lives that may influence game outcomes (Cardenas and Ostrom, 2004). For instance, individuals may believe that the intensity of their own fishing has far fewer consequences on the future availability of fish than do, say, climate factors or other industries. Punishment may fail to increase cooperation also because there is insufficient coordination or no venue for discussing the reasons for punishment. Janssen et al. (2010) argue that, when participants can "punish back" but cannot discuss *why* they are punished, being sanctioned does not carry a clear message. In our experiment, the occurrence of antisocial punishment may be attributable not only to the lack of a unique social norm but also to the lack of discussion about sanctions. Hence punishment failed to transmit the moral lesson that high extraction levels should be sanctioned. Furthermore, those who were punished for extracting *low* levels reacted by increasing their extraction levels. Beckenkamp and Ostmann (1999) and Masclet et al. (2003) argue that punishment will likely reduce cooperation if subjects perceive the sanctions to be unfair (antisocial). It is interesting that subjects were willing to incur a monetary cost in order to administer nonmonetary punishment, a finding that accords with many previous studies on monetary punishment. More generally, the subjects themselves may not expect punishment to induce more cooperative behavior. Carpenter et al. (2004), Carpenter (2007), Casari and Luini (2009), Fudenberg and Pathak (2010), and Noussair et al. (2011) conclude that punishment need not be applied instrumentally to increase cooperation and that subjects have preferences for punishing.

Third, in line with Masclet et al. (2003), Velez et al. (2009) and Hayo and Vollan (2012), we find strong conformity effects: individuals adjust their period-by-period decisions in order to more nearly match their peers' average in the previous period. Also, even cooperative individuals increase their own extraction levels after observing that their current game partners were punished in the previous period. These results indicate that social preferences are context dependent and also suggest that social comparisons may serve as a nonpecuniary way for policy to encourage changes in behavior (Ferraro and Price, 2011).

Fourth, and contrary to most previous research on this topic, we find no in-group bias with respect to cooperation. That is, individuals do not behave differently when interacting with those from their own community than when mixed with subjects from other communities—except for being more sensitive in mixed groups to the threat of NMP. Many fishermen who were surveyed complained about fishers from other communities who arrive during the peak season to fish in the lagoon where the complainants fish year round. Yet fishers from all communities fish also in other locations during seasonal peaks. That our experiment revealed no in-group bias might therefore reflect fishers acknowledging that all of them are sometimes "outsiders". It could also reflect the lack of any meaningful differences among our focal communities in terms of ethnicity or religion, differences that other studies have identified as determinants of negative feelings about out-group members (Hewstone et al., 2002).

Overall our results show that, in response to the threat of nonmonetary punishment, individuals cooperate (here, by limiting their resource exploitation) when they are mixed with individuals from other communities. The implication is that cooperation could be enhanced by incorporating social sanctions into the management of common pool resources that are exploited by more than one community. However, our findings also indicate that coordination is required to render peer punishment effective, to prevent antisocial targeting, and to enhance the social signal conveyed by such punishment. Finally, we establish that previous interactions with other subjects—even if only in a series of one-shot games—exert substantial influence on behavior and so reflect strong conformity preferences.

Acknowledgments

We are particularly thankful to Sam Bowles, Francisco Alpizar and Fredik Carlsson for their invaluable advice. We are also very thankful for the comments received from Gabriele Camera, Luigi Luini, Shachar Kariv, Jordi Brandts, Jorge Maldonado, Jeroen van der Bergh Josep Lluís Raymond and Andres Rius. We would also like to thank comments from the participants of the

LACEEP's XII and XIII workshops. 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 implemented without the commitment and support from researchers that helped us during the experimental sessions. 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., Bardhan, P. and Bowles, S. (2007) *"Inequality, cooperation and environmental sustainability"*. Russell Sage Foundation and Princeton University Press.

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

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

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

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

Bowles, S. and Polanía-Reyes, S. (2012) "Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature*, 50:2, 368-425.

Brandts, J. and Charness, G. (2011) "The strategy versus the direct-response method: a survey of experimental comparisons", *Experimental Economics*, Vol. 14, pp. 375 – 398.

CAEPALR (2006) “Propuesta de ingreso del área protegida laguna de Rocha al sistema nacional de áreas protegidas”, Comisión Asesora Específica Provisoria del Área Protegida Laguna de Rocha.

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., and Ostrom, E. (2004) “What do people bring into the game? Experiments in the field about cooperation in the commons”, *Agricultural Systems*, Vol. 82, pp. 307–326.

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

Charness, G., Rigotti, L. and Rustichini, A. (2007) “Individual Behavior and Group Membership”, *The American Economic Review*, Vol. 97, Nº 4, pp. 1340 – 1352.

Carpenter, J., Matthews, P., Ong’ong’a, O. (2004) “Why punish? Social reciprocity and the enforcement of prosocial norms”, *Journal of Evolutionary Economics*, 14: 407-429.

Carpenter, J. (2007) “The demand for punishment”, *Journal of Economic Behavior and Organization*, Vol. 62: 522-542.

Chen, Y and Xin S. L. (2009) “Group identity and social preferences”. *American Economic Review*, 99(1): 431–57.

Defeo, O., Horta, S., Carranza, A., Lercari, D., de Álava, A., Gómez, J., Martínez, G., Lozoya, J.P. and Celenano, E. (2009) Hacia un manejo ecosistémico de pesquerías: Áreas Marinas Protegidas en Uruguay, DINARA and Facultad de Ciencias – UdelaR.

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. (2012) “The impact of group membership on cooperation and norm enforcement: evidence using random assignment to real social groups”. *American Economic Journal: Microeconomics,* 4:101-115.

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

Hargreaves Heap, S. P., and Zizzo, D.J. (2009) “The Value of Groups”, *American Economic Review*, Vol. 99, Nº 1, pp. 295–323.

Harris, D., Herrmann, B. and Kontoleonx, A. (2012) “When to Favour Your Own group?The Threats of Costly Punishments and In-group Favouritism”, Discussion Paper Series, Nº 628, Department of Economics, University of Oxford.

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

Hayo, B., Vollan, B. (2012) “Group interaction, heterogeneity, rules, and co-operative behaviour: Evidence from a common-pool resource experiment in South Africa and Namibia”, *Journal of Economic Behavior & Organization*, Vol. 81, pp. 9– 28.

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 Gächter, 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.

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

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

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. y 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.

Tajfel, H.; Turner, J. (1979). "An Integrative Theory of Intergroup Conflict", in Austin, W. G.; Worchel, S.. The Social Psychology of Intergroup Relations. Monterey, CA:

Turner, J. C., Brown, R. J., and Tajfel, H. " Social comparison and group interest in ingroup favouritism" *European Journal of Social Psychology*, Vol. 9, pp. 187-204

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: Average socioeconomic characteristics by community

Community	Years of schooling	Electricity at home	Wealth[a]	Per capita income (US)	Fishing main activity
Laguna de Rocha	6.0	13%	1.85	149	75%
Valizas Puente	6.7	75%	3.06	175	67%
Barra de Valizas	7.6	38%	1.68	373	63%
Puerto los Botes	6.0	100%	2.52	246	100%
Barrio Parque	8.0	100%	4.32	320	38%

[a] The wealth index considers different durable goods a household may own.

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: Punishing 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: Net variations, flags, and deviations from group's average in previous round including interactions terms

	Dependent variable: $fishnets_t$ - $fishnets_{t-1}$	
Sample	**7-10 & 17-20**	**7-10 & 17-20**
Positive deviation from average ($a_{i,t-1}$-$\bar{a}_{t-1}$)	-0.952***	-0.940***
	(0.183)	(0.201)
Negative deviation from average ($\bar{a}_{t-1}$-ai,$_{t-1}$)	0.876***	0.934***
	(0.166)	(0.162)
Positive deviation from average ($a_{i,t-1}$-$\bar{a}_{t-1}$)*$flag_{t-1}$	-0.010	
	(0.204)	
Negative deviation from average ($\bar{a}_{t-1}$-$a_{i,t-1}$)*flagt-1	-0.095	
	(0.168)	
Positive deviation from average ($a_{i,t-1}$-$\bar{a}_{t-1}$)*outgroup		-0.076
		(0.166)
Negative deviation from average ($\bar{a}_{t-1}$-$a_{i,t-1}$)*outgroup		-0.235
		(0.156)
_cons	0.120	0.144
	(0.165)	(0.169)
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		