Universidad del Desarrollo
Facultad de Gobierno

EXPERIMENTAL CLASSIFICATION AND COMPUTATIONAL SIMULATION OF STRATEGIC HETEROGENEITY IN COMMON-POOL RESOURCE SOCIAL DILEMMAS

POR: CARLOS ANDRÉS HIDALGO GARRIDO

Tesis presentada a la Facultad de Gobierno de la Universidad del Desarrollo para optar al título de doctor en Ciencias de la Complejidad Social

PROFESOR GUÍA:

Sr. CARLOS ANDRÉS RODRÍGUEZ-SICKERT

CO-GUÍA:

RICARDO ANDRÉS GUZMÁN PRICE

Octubre 2024

SANTIAGO, CHILE

ABSTRACT: This study investigates the role of strategic heterogeneity in the governance of Common-Pool Resources (CPRs), particularly in the context of co-management policies. While effective in preventing the tragedy of the commons, co-management outcomes vary across communities. We propose that differences in co-management performance can be attributed to the heterogeneity of strategic types within user groups and how these distributions shift in response to external enforcement. To explore this, we conducted a repeated CPR game experiment with small-scale fishing communities in Chile, categorized by their real-life experience with co-management (ranging from no experience to high and lower performance). The classification was based on both secondary and primary data, which combined biological, economic, and organizational indicators. Participants were subjected to two treatments: one without norm enforcement and one with non-deterrent external enforcement. We classified their cooperative strategies as free-riders, conditional cooperators, unconditional cooperators, or negative cooperators and analyzed the distribution across user groups. Our findings demonstrate a strong association between strategic heterogeneity and co-management outcomes, highlighting the influence of co-management institutions on shaping cooperative dynamics. Additionally, we used these experimentally informed results to develop an agent-based model that simulates how varying proportions of each strategic behavior affect the stability or erosion of cooperation in social dilemmas. This research underscores the significance of considering strategic heterogeneity and incentive structures in designing effective co-

management interventions, offering deeper insights into the behavioral mechanisms that drive cooperation in CPR management.

1. INTRODUCTION

Our biological, social, and cultural evolution has greatly depended on our ability to cooperate, particularly when confronting social dilemmas that require collective action (Bowles & Gintis, 2013; Henrich et al., 2004). A social dilemma is a scenario in which individual and collective interests are at odds: the optimal collective outcome is achieved when everyone cooperates, yet there are incentives for individuals to maximize their short-term payoffs, thereby jeopardizing long-term collective benefits (Kollock, 1998; Van Lange et al., 2013). Well-known examples of social dilemmas include the provision of public goods and the exploitation of common-pool resources (CPRs). In such scenarios, canonical models of self-interested individuals predict that free-riding (i.e., defecting from cooperation) will emerge as the dominant strategy, restricting the possibility of cooperative behavior. However, extensive evidence from experimental economics and observational case studies reveal that this prediction does not hold, as people tend to cooperate far more than expected (Ostrom, 1990): public goods are frequently created, and the tragedy of the commons can effectively be averted if the right conditions are met. But cooperation in social dilemmas is also complex and fragile. In repeated interactions without incentives such as rewards or punishment, cooperation levels often exhibit a consistent downward spiral over time (Ledyard, 1995; Spadaro et al., 2022).

The sustainable management of CPRs such as fisheries, forestry, irrigation systems, and pastures, remains a critical challenge in environmental and resource economics (Nogueira et al., 2021; Schill et al., 2016; Tavoni & Levin, 2014). Central to addressing this challenge is a deep understanding of the decision-making processes of CPR users, which is crucial for crafting effective management strategies tailored to complex environmental issues (Blanco & Walker, 2019; Carattini et al., 2019; Kluvankova et al., 2019; Rodela et al., 2019; Tucker, 2010). In recent decades, co-management has gained attention as a policy mechanism to address the threat of the tragedy of the commons. Co-management is a formal institution implemented by States to strengthen CPR and other natural resources governance, and its main characteristic is that it assigns joint-responsibility in the design and enforcement of management plans among state agencies, CPR users, and scientists. Scientific assessments of the effectiveness of co-management as an institutional arrangement for the sustainable exploitation of CPRs have not fully explored how or why the same set of rules differentially impacts small-scale CPR user groups that share a cultural background, geographic region, and ecological characteristics.

In the current state of the field significant attention has been directed toward the proximate roles of institutional and psychological variables influencing cooperation and its stability. This thesis addresses two of these mechanisms and analyzes the interaction between them as a potential explanation for the varying levels and dynamics of cooperation in the co-management of common-pool resources: the enforcement of social norms through external punishment and the role of strategic heterogeneity.

Studies of norm enforcement through incentives (rewards or punishment) have a long-standing tradition in experimental economics (Anderson & Stafford, 2003; Balliet et al., 2011; Balliet & Van Lange, 2013; Boyd et al., 2010; Fehr & Gächter, 2000a; Gächter et al., 2008). However, the literature on strategic heterogeneity and its interaction with these incentives is considerably smaller. Much of the work on heterogeneity has focused on preferences, particularly since the early theoretical and experimental studies on other-regarding preferences (Bolton, 1991; Bolton & Ockenfels, 2000; Camerer & Thaler, 1995; Charness & Rabin, 2002; Cooper & Kagel, 2016; Dawes & Thaler, 1988; Falk et al., 2003; Fehr & Schmidt, 1999; Sobel, 2005). In essence, other-regarding preferences refer to individual concerns not only for one's own well-being but also for the well-being of others. These preferences have been proposed as a key explanation for the higher levels of cooperation observed, which exceed the predictions of classical economic and evolutionary models based on self-regarding preferences.

One of the premises of this thesis is that preferences are not the same as strategies (Velez et al., 2009). By strategies we refer to complete plans of action in response to the cooperative decisions of others and a given incentive structure. These plans of action may align with either self-interested or social preferences, such as altruism, reciprocity, or inequity aversion, thereby embodying varying degrees of cooperation. Strategic heterogeneity has been consistently observed through experimental game theory, leading to the characterization of strategies such as conditional cooperation, unconditional cooperation, and free-riding, among others (Fallucchi et al., 2019; Fischbacher et al., 2001; Kurzban & Houser, 2005). Strategies, therefore, are interdependent responses that take into account preferences but also react to different group-level characteristics and structural incentives (Cárdenas & Ostrom, 2004).

In many real-world settings, CPR user groups address social dilemmas by enforcing a cooperative social norm established in a co-management plan. However, enforcement mechanisms are almost always imperfect, with punitive measures that are either too low or uncertain. When confronted with these non-deterrent enforcement mechanisms, CPR users may adjust (or resist adjusting) their strategies, likely influenced by changes in their expectations about the group's adherence to the social norm. The research questions we seek to answer is how CPR users respond -whether by adapting or maintaining their strategies - when a non-deterrent norm enforcement incentive is introduced, and how varying proportions of these strategies explain the group's cooperative dynamics both in the absence of external enforcement and after the enforcement is in place. From this research question, we derive two main objectives: 1) to classify individual strategies across different CPR samples under varying incentive structures to assess group dynamics as an outcome of their strategic profiles; and 2) to explore the effect of the prevalence of each strategy within different CPR user group samples.

We address these objectives separately in the two research papers presented below. In Paper 1 (“Shifting strategies: exploring cooperation dynamics in fisheries co-management”, currently under peer-review in Ecology & Society), we explain the pre-classification of the Chilean fishers’ sample categorized into user types of CPR user groups based on their real-life experience with co-management (no experience, highly performing, and lower performance) (Gelcich et al., 2013), and that to control for the large biological and sociocultural variability, the subject pool was selected from a limited area of the central coast of Chile. We then adapt a method known as the Linear Conditional Contribution Profile to classify individual strategies in each treatment and estimate the strategic profile (i.e., the proportions of each strategy) for each type of CPR user group. Finally, we associate the real-life performance of each user group with their strategic profile and discuss what might have motivated different strategic responses to the non-deterrent treatment. In Paper 2 (“Strategic heterogeneity and cooperation dynamics in common-pool resource dilemmas: experimental and simulation insights”, currently ready for peer-review), we design an agent-based model informed by the experimental results analyzed in Paper 1 to test the effect of the prevalence of each strategy across different types of CPR samples. We estimate the linear function of each strategy and use it to simulate the effect of their increasing prevalence (from 0 to 100%) on the stability of cooperation within different types of CPR groups.

Both papers use the experimental results of a lab-in-the-field contextually framed repeated CPR game conducted with 85 small-scale fishers from the central coast of Chile. In the experiment, all participants played the CPR game in two treatments: one without enforcement of a social norm and one with non-deterrent external enforcement of a social norm. We then classified fishers' individual cooperative strategies in each treatment as either free-riders, conditional cooperators, unconditional cooperators, or negative cooperators, and assessed the distribution of strategies across types of user groups in both treatments.

Our contributions to the field of social dilemmas and CPR co-management are threefold. First, we bridge two strands of literature that have largely remained disconnected. As previously mentioned, there is limited experimental evidence on the interaction between heterogeneity and norm enforcement, and virtually no evidence on how norm enforcement affects individuals' strategy-switching behavior (with the notable exception of Rodriguez-Sickert et al., 2008, although their work is more closely aligned with the literature on preferences). Moreover, zooming into the individual and group levels allows us to incorporate new insights from behavioral and experimental economics into our understanding of co-management. Although experimental game theory has largely explored the role of norm enforcement in cooperative dynamics, the literature on co-management has rarely applied this approach to address the effectiveness of these institutions nor explored how or why the same set of rules differentially impacts small-scale

CPR user groups that share a cultural background, geographic region, and ecological characteristics. Secondly, we address this gap by employing a contextually framed lab-in-the-field experiment with a pre-classified sample, a key feature that enhances the internal, ecological, and external validity of our results. Furthermore, it is essential to highlight the importance of studying natural populations in the context of CPRs management and cooperation dynamics. Natural populations (such as CPR users) provide a unique opportunity to observe real-world interactions and the influence of institutional arrangements, cultural norms, and ecological conditions on cooperative behavior. Chile, with its extensive coastline and diverse fishing communities, presents an ideal case for examining how CPR users navigate the challenges of resource management. The country's co-management policies, particularly within the context of small-scale fisheries, offer valuable insights into the interplay between natural populations, strategic heterogeneity, and institutional frameworks. Finally, through the design of an experimentally informed agent-based model, we contribute counterfactual results that explore a broader range of possible outcomes and more robustly test the effects of each strategy on potential real-world CPR scenarios. Thirdly, this research is further contributing to understanding how changes in normative expectations influence individual decision-making and strategy adaptation, which is essential for developing more effective, context-sensitive approaches to CPR co-management. In this sense, our study provides new insights into the co-management literature by introducing methods form experimental economics into the more specific field of co-management.

This document is organized as follows: in chapter 2 we provide a general overview of the evidence on punishment and strategic heterogeneity in cooperative behavior within social dilemmas. Chapter 3 presents the research paper addressing the first objective of this thesis: the classification of individual strategies and the estimation of strategy prevalence across different types of CPR user groups. Finally, Chapter 4 presents the research paper related to our second objective: the design of an agent-based model to estimate the cooperative outcomes of varying strategy prevalence within real-world strategy distributions estimated in our Chilean CPR sample.

2. RELATED LITERATURE

2.1. NORM ENFORCEMENT BY PUNISHMENT IN EXPERIMENTAL SOCIAL DILEMMAS

The existence and extent of pro-social behaviors in social dilemmas present a puzzle to scholars and has prompted a significant focus on studying norm enforcement mechanisms. Norms are commonly understood as widely shared standards of behavior that dictate how to act in specific situations (Fehr & Schurtenberger, 2018) and in the context of social dilemmas they specify how, when, and with whom to cooperate. Individuals regularly face situations where these norms significantly shape their decision-making, and they often use various mechanisms to enforce compliance, ranging from direct punishment (e.g., confrontation) to indirect punishment (e.g., gossip and social exclusion) (Balliet et al., 2022; Molho et al., 2024). The experimental literature on social dilemmas has thoroughly examined the impact of norm enforcement through punishment, whether via economic or social sanctions, with a particular focus on how and punishing free-riders or even low cooperators can enhance collective benefits (Anderson & Stafford, 2003; Gächter et al., 2008; Gürerk et al., 2006; Shinada & Yamagishi, 2007; Yamagishi, 1986, 1988). In a meta-analysis of the effects of incentives such as rewards and punishments on cooperation, Balliet et al. (2011) analyzed 76 papers and 187 effect sizes from scientific reports on social dilemmas. After controlling for experimental design variables, the authors found a medium-sized effect of punishment on cooperation and evidence that punishment is more effective in repeated games than in one-shot games. Additionally, they observed that decentralized punishment (administered by participants in the social dilemma) was more common in experiments (k=102) than centralized punishment treatments (administered by an external source, such as the experimenter: k=83). In a more recent meta-analysis of cross-cultural cooperation, Spadaro et al. (2022) tested 17 mainstream hypotheses commonly used to explain cooperation across 1,506 studies conducted in 70 societies. Using an econometric approach with multi-level regression models, the authors found that only a few variables positively correlate with cooperation across the experimental approaches analyzed, including the presence of communication (versus no communication), the number of choice options in the experimental design, and, most importantly for our research, the presence of punishment. The role of punishment is significant not only across the entire meta-analysis sample but also in explaining cross-cultural variations in cooperation, suggesting differing cultural sensitivities to it (Herrmann et al., 2008).

The effects of punishment on cooperations are, however, complex. For instance, costly peer-punishment has found experimental evidence of being successful to maintain cooperation stable (Fehr & Gächter, 2000), although its efficacy depends on the costs for punishers, the time span modeled by the experimenters, the rate between pro-social punishment (punishment of free-riders by cooperators) and anti-social punishment (punishment of cooperators by free-riders), and solving the second-order dilemma of punishing subjects who don't punish free riders, among other conditions (Chaudhuri, 2011; Herrmann et al., 2008, 2008; Mathew, 2017; Rand et al., 2010). Recent evidence from CPR games also highlights a similar overall effectiveness of decentralized (peer-to-peer) versus centralized (external or top-down) punishment mechanisms, even though decentralized mechanisms appear to be more efficient in disciplining heavy over-exploiters (Vollan, 2008). Punishment can also produce a crowding-out effect, altering intrinsic motivations to cooperate and leading to a detrimental long-term impact on collective action (Cardenas et al., 2004; Gneezy & Rustichini, 2000; Vollan, 2008). Alternatively, it can generate crowding-in effects, where individuals adjust their strategies toward cooperation (Bowles & Polania-Reyes, 2012; Rodriguez-Sickert et al., 2008).

For the purposes of this study, a key feature of punishment is that it might operate through social norm expectations. Deterrent punishment refers to an enforcement mechanism that alters the incentive structure in such a way that cooperation becomes the dominant strategy. However, in real life, many punishment mechanisms are imperfect, either due to low sanctions or uncertainty in their implementation. In such cases, this is referred to as non-deterrent punishment, where cooperation becomes individually beneficial only if one expects others to adhere to the norm. In our study, we use an external economic non-deterrent punishment for two reasons: first, it provides an ecologically valid representation of the type of enforcement faced by the subjects in our sample, who experience this form of imperfect enforcement from state authorities overseeing their co-management areas. Second, it presents a more compelling alternative to deterrent enforcement, as it does not make cooperation the dominant strategy but instead activates preferences and expectations by signaling a normative standard (Bicchieri et al., 2021; Fehr & Schurtenberger, 2018). On a topic that is particularly important for our purpose, Shinada & Yamagishi (2007) the possibility of direct and indirect effects of punishment on pro-social behavior. They found that punishment triggers a direct effect when it changes the outcomes of cooperation and defection in such a way that renders cooperation more profitable than defection, but, in addition, it also triggers an indirect effect on the expectation of conditional cooperators. This indirect effect provides on conditional cooperators a higher expectation that other subjects will cooperate, since the threat of punishment will enforce other subjects to avoid free-riding, and, just as important, these effects are complementary, as the net effect of punishment can be disaggregated on both direct and indirect.

2.2. HETEROGENEITY OF PREFERENCES AND STRATEGIES IN COOPERATIVE BEHAVIOR

As previously mentioned, another explanation for the extent of human cooperation in social dilemmas is preference heterogeneity. Both laboratory and field experiments have provided evidence for the existence other-regarding preferences that can account for the observed cooperative behaviors (Bolton &

Ockenfels, 2000; Fehr & Schmidt, 1999; Poulsen & Poulsen, 2006). Although the ultimate causes behind the heterogeneity of preferences remain a subject of debate, there is strong evidence supporting preference heterogeneity, which in turn facilitates the classification of individuals—a topic that has garnered significant attention since the introduction of the Strategy Method. In a prominent study, Fischbacher et al. (2001) applied the Strategy Method in an experiment designed to explicitly link individual contributions to each possible average group contribution in a one-shot public goods game, effectively capturing individuals' contributions as a function of other participants' contributions. The original results of Fischbacher et al. estimated that approximately more than half of the subjects behave as conditional cooperators (individuals who cooperate while others also cooperate and who cooperate as much, or nearly as much, as others do), approximately a third part of the sample behave as a free-rider (subjects who do not cooperate regardless of others contributions), and the remaining subjects behave as hump-shaped cooperators (with conditional cooperation until a threshold on others' compliance, and, since then, with declining contributions) or as unclassifiable. In their work, conditional cooperation is a type of reciprocity concern but an imperfect one, as it shows a "self-serving bias", meaning that individuals tend to contribute slightly less than the observed average contribution of the group (Fischbacher & Gächter, 2010).

Since its publication, the Strategy Method has been widely replicated and adapted to analyze the role of conditional cooperation in social dilemmas. For example, Thöni & Volk (2018) examined 17 replications of the original Strategy Method to refine the classification provided by Fischbacher, Gächter, and Fehr (2001). Their work ultimately confirmed the proportionality of free-riders and conditional cooperators found in the original results. Additionally, the Strategy Method has been applied to analyze the ubiquity of conditional cooperation across cultures. For instance, Kocher et al. (2008) provided evidence supporting the hypothesis of widespread cross-cultural behavior with high variance between nations, and Cherry et al. (2017) further explored its correlation with cultural worldviews. These findings help explain why repeated social-dilemma experiments often show more cooperation than expected, such as positive voluntary contributions in public goods games (Chaudhuri, 2011; Ledyard, 1995), but also why cooperation erodes over time, since the interaction of free-riders and conditional cooperators will induce a decrease on the contributions of the latter (Ambrus & Pathak, 2011). Moreover, even in the absence of free-riders, conditional cooperation alone will cause a decline in average contributions due to the self-serving bias as a weighted belief on other’s contributions and self-beliefs on previous rounds.

Another branch of the literature has focused on the effects of group composition and formation mechanisms on cooperation in social dilemmas. Gächter & Thöni (2005) tested whether cooperation among like-minded individuals differs from cooperation in randomly composed groups, with or without peer-punishment. Their results suggest that like-minded groups sustain higher cooperation levels, especially when enforcement is not present. Similarly, Burlando & Guala (2005) classified subjects into homogeneous groups of free-riders, cooperators, and reciprocators and found that homogeneous groups, particularly of cooperators and reciprocators, maintained high and stable contributions. In contrast, groups of free-riders saw contributions collapse by the end of the game. Lastly, de Oliveira et al. (2015) shed light on the interactions between cooperative preferences by examining how group composition affects contributions in voluntary contribution mechanisms (VCM). Their results showed that the presence of free-riders reduces overall cooperation and induces lower contributions from conditional cooperators, highlighting the detrimental impact of selfish behavior on group dynamics.

Other studies have also incorporated norm enforcement treatments (like external or peer punishment) into the preference heterogeneity analysis. Ones & Putterman (2007) incorporated both contributions and punishment behaviors to classify and analyze preference heterogeneity and group performance. They found that knowing the composition of a group allows for accurate predictions of group outcomes and that different assembling criteria can enhance efficient outcomes depending on the problem at hand. Cheung (2014) adapted the Strategy Method to include a punishment treatment and found that, without punishment, conditional contributions decline over time not only due to self-serving bias but also when the distribution of others' contributions becomes more unequal. In an important contribution, Rustagi et al. (2010) combined both experimental and observational field data to investigate the role of conditional cooperation and punishment in the sustainable exploitation of forest common-pool resources. Their results from a one-shot public goods game using the Strategy Method revealed that groups with a higher proportion of conditional cooperators were more successful in managing forest commons. This success was largely attributed to the fact that individuals with this preference type were more inclined to incur the cost of punishing free-riders and enforcing cooperation.

Using a different approach to classify preference types, Kurzban & Houser (2005) developed the Linear Conditional Contribution Profile (LCCP) to analyze the results of a repeated voluntary contribution game. The LCCP (which we also use in our research) fits individual responses to the mean observed cooperation by other subjects in the previous round using a linear model and uses the parameters of this model to classify subjects' strategies. Contrary to the Strategy Method (more related to the identification of preferences), the LCCP estimates only consistent plans of action, which resembles more to our definition of strategies rather than preferences. Kurzban & Houser (2005) found significative differences between cooperative strategies and, more importantly, evidence of preference stability across rounds. However, a more recent result challenge this results and highlights that, despite preference stability reported before, conditional cooperators can transform into free-riders and that this process is partially explained by their beliefs of other subject’s choices, although the opposite transformation is seldom observed (Andreozzi et al., 2020). It is noteworthy that the LCCP has demonstrated a distribution of cooperative strategies similar to the preference distribution observed through the Strategy Method, providing strong evidence of the accuracy of this approach (van Klingeren, 2022).

Three studies are of particular interest due to their methodological and/or research question, as they are closely related to this thesis. Rodriguez-Sickert et al. (2008) used experimental data to estimate preferences and analyze the effect of external enforcement on this preference distribution (like our research question). They conducted a repeated common-pool resource game with 320 Colombian villagers and tested four treatments: a No Fine institution (NF), a High Fine institution (HF), a Low Fine institution (LF), and a proposed but rejected fine institution (RF). Their results highlighted several key findings relevant to our research: fines, even low ones, can stabilize cooperation; external enforcement can alter the distribution of preferences; and that external enforcement affects behavior through deterrence and moralization (with the latter associated with the LF treatment). In their results, the moralization effect on behavior accounts for most of the changes in preferences in both HF and LF treatments, which explains why, in their experiments, both institutions have quiet similar positive effect on stabilizing cooperation across rounds. Our work can be understood as an extension of Rodriguez-Sickert et al. (2008), as we try to respond to a similar question, although our contribution is to use a pre-classified sample which enables us to a more detailed data set to analyze how individuals from groups with different real-life performances on co-management react to the norm enforcement effect of non-deterrent external punishment.

The other two studies are relevant to our research due to their methodological similarities to our approach, particularly in the methods used to classify strategic types and assess strategic profiles within groups. Kurzban & Houser (2005) use a repeated voluntary contribution mechanism (VCM) game (i.e., a repeated public goods game) and proposes a linear model to classify subjects as either free-riders, conditional cooperators or cooperators (the latter being similar to our unconditional cooperators). Their classification method utilized each subject's Linear Conditional Contribution Profile (LCCP), estimated through an OLS regression of their contribution on the mean contribution observed immediately before contributing. According to Kurzban and colleagues, the results of the LCCP provide, for each subject, an intercept parameter indicating the willingness to cooperate when teammates contribute nothing or very little to the public good, and a slope parameter representing the extent of conditional cooperation. Therefore, the authors were able to estimate that nearly 20% of the sample could be classified as free-riders, 63% as reciprocators (conditional cooperators) and 13% as cooperators. who also used a repeated VCM that included a treatment allowing subjects to express disapproval of their teammates' contributions. To classify social preferences, they regressed each subject's contribution on the contributions of their teammates in the previous round, thereby estimating a predisposition toward unconditional cooperation (the intercept of the model) and a predisposition toward conditional cooperation (the slope). Using this method, the authors were able to estimate that fishermen who jointly engaged in economic activities by sharing expenses and income (referred to as poolers) were less likely to be conditional cooperators and more likely to be unconditional cooperators compared to non-poolers. Additionally, crews with higher levels of conditional cooperation and a greater propensity to disapprove of defections were found to be more productive. As will be detailed in the methodological section of Paper 1, our classification method employs a very similar approach to those used by Kurzban & Houser and Carpenter & Seki, but we applied it to address a research question more akin to that posed by Rodriguez et al.

All this evidence can be summarized in the following stylized facts about heterogeneous preferences in social dilemmas: i) individuals can be successfully categorized into discrete or continuous types of cooperativeness; ii) conditional cooperators account for slightly more than half of the individuals and free-riders for approximately a third of the proportion of individuals in a given group; iii) knowing the proportion of cooperation types can successfully predict group performance; iv) although preferences are seemingly stable across time, under certain circumstances (such as external enforcement of norms), subjects can adapt their behavior and act according to a different preference type.

2.3. REFERENCES

Ambrus, A., & Pathak, P. A. (2011). Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics*, *95*(7), 500-512. https://doi.org/10.1016/j.jpubeco.2010.11.016

Anderson, L. R., & Stafford, S. L. (2003). Punishment in a Regulatory Setting: Experimental Evidence from the VCM. *Journal of Regulatory Economics*, *24*(1), 91-110. https://doi.org/10.1023/A:1023952115422

Andreozzi, L., Ploner, M., & Saral, A. S. (2020). The stability of conditional cooperation: Beliefs alone cannot explain the decline of cooperation in social dilemmas. *Scientific Reports*, *10*(1), 13610. https://doi.org/10.1038/s41598-020-70681-z

Balliet, D., Molho, C., Columbus, S., & Dores Cruz, T. D. (2022). Prosocial and punishment behaviors in everyday life. *Current Opinion in Psychology*, *43*, 278-283. https://doi.org/10.1016/j.copsyc.2021.08.015

Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *137*(4), 594-615. https://doi.org/10.1037/a0023489

Bicchieri, C., Dimant, E., & Xiao, E. (2021). Deviant or wrong? The effects of norm information on the efficacy of punishment. *Journal of Economic Behavior & Organization*, *188*, 209-235. https://doi.org/10.1016/j.jebo.2021.04.002

Blanco, E., & Walker, J. M. (2019). Common-pool resource appropriation and conservation: Lessons from experimental economics. En *Routledge Handbook of the Study of the Commons*. Routledge.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, *90*(1), 166-193. https://doi.org/10.1257/aer.90.1.166

Bowles, S., & Gintis, H. (2013). *A Cooperative Species: Human Reciprocity and Its Evolution*.

Bowles, S., & Polania-Reyes, S. (2012). Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature*, *50*(2), 368-425. https://doi.org/10.1257/jel.50.2.368

Burlando, R. M., & Guala, F. (2005). Heterogeneous Agents in Public Goods Experiments. *Experimental Economics*, *8*(1), 35-54. https://doi.org/10.1007/s10683-005-0436-4

Carattini, S., Levin, S., & Tavoni, A. (2019). Cooperation in the Climate Commons. *Review of Environmental Economics and Policy*, *13*(2), 227-247. https://doi.org/10.1093/reep/rez009

Cardenas, J.-C., Ahn, T. K., & Ostrom, E. (2004). Communication and Co-operation in a Common-Pool Resource Dilemma: A Field Experiment. En S. Huck (Ed.), *Advances in Understanding Strategic Behaviour: Game Theory, Experiments and Bounded Rationality* (pp. 258-286). Palgrave Macmillan UK. https://doi.org/10.1057/9780230523371_12

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, *14*(1), 47-83. https://doi.org/10.1007/s10683-010-9257-1

Cherry, T. L., McEvoy, D. M., & Sælen, H. (2017). Conditional cooperation and cultural worldviews. *Economics Letters*, *158*, 51-53. https://doi.org/10.1016/j.econlet.2017.06.034

Cheung, S. L. (2014). New insights into conditional cooperation and punishment from a strategy method experiment. *Experimental Economics*, *17*(1), 129-153. https://doi.org/10.1007/s10683-013-9360-1

de Oliveira, A. C. M., Croson, R. T. A., & Eckel, C. (2015). One bad apple? Heterogeneity and information in public good provision. *Experimental Economics*, *18*(1), 116-135. https://doi.org/10.1007/s10683-014-9412-1

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *The American Economic Review*, *90*(4), 980-994.

Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, *114*(3), 817-868.

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, *2*(7), 458-468. https://doi.org/10.1038/s41562-018-0385-5

Fischbacher, U., & Gächter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review*, *100*(1), 541-556. https://doi.org/10.1257/aer.100.1.541

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397-404. https://doi.org/10.1016/S0165-1765(01)00394-9

Gächter, S., Renner, E., & Sefton, M. (2008). The Long-Run Benefits of Punishment. *Science*, *322*(5907), 1510-1510. https://doi.org/10.1126/science.1164744

Gächter, S., & Thöni, C. (2005). Social Learning and Voluntary Cooperation Among Like-Minded People. *Journal of the European Economic Association*, *3*(2/3), 303-314.

Gelcich, S., Guzman, R., Rodríguez-Sickert, C., Castilla, J. C., & Cárdenas, J. C. (2013). Exploring External Validity of Common Pool Resource Experiments: Insights from Artisanal Benthic Fisheries in Chile. *Ecology and Society*, *18*(3). https://doi.org/10.5751/ES-05598-180302

Gneezy, U., & Rustichini, A. (2000). A Fine Is a Price. *The Journal of Legal Studies*, *29*(1), 1-17. https://doi.org/10.1086/468061

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The Competitive Advantage of Sanctioning Institutions. *Science*, *312*(5770), 108-111. https://doi.org/10.1126/science.1123633

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (2004). *Foundations of Human Sociality*. Oxford University Press. https://doi.org/10.1093/0199262055.001.0001

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial Punishment Across Societies. *Science*, *319*(5868), 1362-1367. https://doi.org/10.1126/science.1153808

Kluvankova, T., Brnkalakova, S., Gezik, V., & Maco, M. (2019). Ecosystem services as commons? En *Routledge Handbook of the Study of the Commons*. Routledge.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., & Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters*, *101*(3), 175-178. https://doi.org/10.1016/j.econlet.2008.07.015

Kollock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, *24*(1), 183-214. https://doi.org/10.1146/annurev.soc.24.1.183

Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences*, *102*(5), 1803-1807. https://doi.org/10.1073/pnas.0408759102

Ledyard, J. O. (1995). *Public Goods: A Survey of Experimental Research* (J. H. Kagel & A. E. Roth, Eds.; pp. 111-194). Princeton University Press. https://resolver.caltech.edu/CaltechAUTHORS:20171107-163442067

Mathew, S. (2017). How the Second-Order Free Rider Problem Is Solved in a Small-Scale Society. *American Economic Review*, *107*(5), 578-581. https://doi.org/10.1257/aer.p20171090

Molho, C., De Petrillo, F., Garfield, Z. H., & Slewe, S. (2024). Cross-societal variation in norm enforcement systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1897), 20230034. https://doi.org/10.1098/rstb.2023.0034

Nogueira, L. A., Wigger, K. A., & Jolly, S. (2021). Common-pool resources and governance in sustainability transitions. *Environmental Innovation and Societal Transitions*, *41*, 35-38. https://doi.org/10.1016/j.eist.2021.10.004

Ones, U., & Putterman, L. (2007). The ecology of collective action: A public goods and sanctions experiment with controlled group formation. *Journal of Economic Behavior & Organization*, *62*(4), 495-521. https://doi.org/10.1016/j.jebo.2005.04.018

Ostrom, E. (1990, noviembre). Governing the Commons: The Evolution of Institutions for Collective Action. En *Cambridge Core*. https://doi.org/10.1017/CBO9780511807763

Poulsen, A., & Poulsen, O. (2006). Endogenous Preferences and Social-Dilemma Institutions. *Journal of Institutional and Theoretical Economics*, *162*(4), 627. https://doi.org/10.1628/093245606779252742

Rand, D. G., Armao IV, J. J., Nakamaru, M., & Ohtsuki, H. (2010). Anti-social punishment can prevent the co-evolution of punishment and cooperation. *Journal of Theoretical Biology*, *265*(4), 624-632. https://doi.org/10.1016/j.jtbi.2010.06.010

Rodela, R., Tucker, C. M., Šmid-Hribar, M., Sigura, M., Bogataj, N., Urbanc, M., & Gunya, A. (2019). Intersections of ecosystem services and common-pool resources literature: An interdisciplinary encounter. *Environmental Science & Policy*, *94*, 72-81. https://doi.org/10.1016/j.envsci.2018.12.021

Rodriguez-Sickert, C., Guzmán, R. A., & Cárdenas, J. C. (2008). Institutions influence preferences: Evidence from a common pool resource experiment. *Journal of Economic Behavior & Organization*, *67*(1), 215-227. https://doi.org/10.1016/j.jebo.2007.06.004

Rustagi, D., Engel, S., & Kosfeld, M. (2010). Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science*, *330*(6006), 961-965. https://doi.org/10.1126/science.1193649

Schill, C., Wijermans, N., Schlüter, M., & Lindahl, T. (2016). Cooperation Is Not Enough—Exploring Social-Ecological Micro-Foundations for Sustainable Common-Pool Resource Use. *PLOS ONE*, *11*(8), e0157796. https://doi.org/10.1371/journal.pone.0157796

Shinada, M., & Yamagishi, T. (2007). Punishing free riders: Direct and indirect promotion of cooperation. *Evolution and Human Behavior*, *28*(5), 330-339. https://doi.org/10.1016/j.evolhumbehav.2007.04.001

Spadaro, G., Graf, C., Jin, S., Arai, S., Inoue, Y., Lieberman, E., Rinderu, M. I., Yuan, M., Van Lissa, C. J., & Balliet, D. (2022). Cross-cultural variation in cooperation: A meta-analysis. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspi0000389

Tavoni, A., & Levin, S. (2014). Managing the climate commons at the nexus of ecology, behaviour and economics. *Nature Climate Change*, *4*(12), 1057-1063. https://doi.org/10.1038/nclimate2375

Thöni, C., & Volk, S. (2018). Conditional cooperation: Review and refinement. *Economics Letters*, *171*, 37-40. https://doi.org/10.1016/j.econlet.2018.06.022

Tucker, C. (2010). Learning on governance in forest ecosystems: Lessons from recent research. *International Journal of the Commons*, *4*(2), 687-706. https://doi.org/10.18352/ijc.224

van Klingeren, F. (2022). Using player types to understand cooperative behaviour under economic and sociocultural heterogeneity in common-pool resources: Evidence from lab experiments and agent-based models. *PloS One*, *17*(5), e0268616. https://doi.org/10.1371/journal.pone.0268616

Van Lange, P. A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125-141. https://doi.org/10.1016/j.obhdp.2012.11.003

Vollan, B. (2008). Socio-ecological explanations for crowding-out effects from economic field experiments in southern Africa. *Ecological Economics*, *67*(4), 560-573. https://doi.org/10.1016/j.ecolecon.2008.01.015

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*(1), 110-116. https://doi.org/10.1037/0022-3514.51.1.110

Yamagishi, T. (1988). Seriousness of Social Dilemmas and the Provision of a Sanctioning System. *Social Psychology Quarterly*, *51*(1), 32-42. https://doi.org/10.2307/2786982

3. FIRST PAPER. SHIFTING STRATEGIES: EXPLORING COOPERATION DYNAMICS IN FISHERIES CO-MANAGEMENT

3.1. INTRODUCTION

Small-scale societies that make use of common-pool resources (CPRs) are exposed to the depletion of the resource system in the so-called "Tragedy of the Commons" (Hardin 1968), a social dilemma in which individual interests collide with collective goals, threatening both the resource system and the communities that make use of it. This potential tragedy endangers a vast array of CPRs and, therefore, has been thoroughly investigated in cases of fisheries (Berkes 1985, Kraak 2011), other marine-coastal resources (Wilkinson and Salvat 2012), grasslands (Cole et al. 2014), irrigation systems (Podimata and Yannopoulos 2015) and forests (Brown and Harris Jr 2008). Case studies and observational evidence have shown that CPRs can undergo processes characterized by diminishing cooperation and escalating resource exploitation under inadequate governance regimes (Ostrom 1990, Castilla and Fernandez 1998, Klooster 2000, Gautam and Shivakoti 2005, Arias Schreiber 2012, Fleischman et al. 2014, Defeo et al. 2016).

In recent decades, co-management has increasingly gained attention as a policy mechanism to address the ongoing threat of the tragedy of the commons. Co-management is a governance approach associated with CPR management that establishes a partnership between public and private actors. It ranges between a simple exchange of information between stakeholders to the joint design and implementation of formal arrangements, combining peer and non-deterrent external enforcement (Berkes et al. 1991, Sen and Raakjaer Nielsen 1996, Plummer and Fitzgibbon 2004, Carlsson and Berkes 2005). Co-management pursues the engagement of different stakeholders, including state agencies, local inhabitants, and users of the resource system (Berkes et al. 1991, Pomeroy and Berkes 1997), while seeking to implement some of the institutional design principles identified for the sustainable management of CPRs, including collective-choice arrangements and a minimal recognition of rights to organize to promote cooperation among users (Ostrom 1990).

Overall, scientific evidence suggests mixed results of co-management policies in a diverse array of CPRs, institutional designs, and sociocultural settings (Evans et al. 2011, Cronkleton et al. 2012, d'Armengol et al. 2018). The literature in this field has primarily concentrated on outcomes that reflect cooperation at the aggregated user group level (Varughese and Ostrom 2001, Tole 2010, Cox et al. 2010, Gutiérrez et al. 2011, Wamukota et al. 2012, Zhu et al. 2014, Defeo et al. 2016, Baggio et al. 2016). While extremely informative about the drivers of co-management variability, this approach does not allow us to delve deeper into the individual motivations and within-group behavioral mechanisms, such as norms and expectations, that affect cooperation dynamics and, consequently co-management outcomes. To better support co-management policies, it is crucial to understand how these groups differ in their tendencies to sustain cooperation and how co-management institutions, such as norm enforcement, interact with their social structures to shape cooperation dynamics.

Sustainability scholars have for long turned to experimental economics to grasp the erosion and stability of cooperation in CPRs (Ostrom 2006, Pisor et al. 2020, Naar 2020). To explain different levels and dynamics of cooperation, two mechanisms have been proposed: strategic heterogeneity and the enforcement of cooperative social norms. Strategies are individual plans of action in response to cooperative decisions by others and to a certain incentive structure. Three strategies have been consistently documented in experimental economics: free-riding (a strategy involving minimal to no cooperation, irrespective of the actions of others), conditional cooperation (the act of cooperating to almost match the observed cooperation of others), and unconditional cooperation (a strategy involving extensive or complete cooperation, regardless of the behavior exhibited by others: Carpenter and Seki, 2011; Cheung, 2014; Fischbacher et al., 2001; Fischbacher and Gächter, 2010; Kocher et al., 2008; Kurzban and Houser, 2005; Rivera-Hechem et al., 2020; Rodriguez-Sickert et al., 2008; Rustagi et al., 2010; van Klingeren, 2022).

Evidence suggests that the erosion of cooperation in experimental repeated social dilemmas (Ledyard 1995, Chaudhuri 2011) is explained by strategic heterogeneity, in particular by the interaction between free-riders and conditional cooperators (as free riders generate a decrease in the contributions of the conditional cooperators), and by the interaction of conditional cooperators with other conditional cooperators, given the imperfect reciprocity or self-serving bias often deployed by conditional cooperators (whose contributions, even in the absence of free-riders, would decrease in repeated interactions: Cheung, 2014; Fischbacher et al., 2001; Fischbacher and Gächter, 2010; Herrmann and Thöni, 2009; Kocher et al., 2008; Thöni and Volk, 2018). While a handful of studies have considered conditional cooperation to explain outcomes in real-life CPR settings (Rustagi et al. 2010, Carpenter and Seki 2011), the role of other cooperative strategies, that could provide a more nuanced understanding of why co-management varies across user groups, has not received the attention it deserves.

In contrast to the role of cooperative strategies, the role of external enforcement in CPRs has been widely studied, both in lab experiments and real-life settings, including co-management arrangements (Cardenas et al. 2000, Travers et al. 2011, Santis and Chávez 2015). In CPR games, the introduction of enforcement generally increases cooperation (Ledyard 1995, Chaudhuri 2011, Fehr and Schurtenberger 2018). However, its effects can be mixed, particularly in field settings. External enforcement has sometimes been shown to crowd out intrinsic motivations to cooperate, negatively affecting long-term collective action (Gneezy and Rustichini 2000, Cardenas et al. 2000 p. 200, Vollan 2008), while in other cases, it crowds in intrinsic incentives to cooperate. These crowding-in and crowding-out effects may reflect shifts in cooperative strategies. Evidence suggests that the introduction of enforcement can modify strategic dispositions toward cooperation (Rodriguez-Sickert et al. 2008, Bowles and Polania-Reyes 2012). When enforcement is deterrent, these strategy shifts often occur due to changes in financial incentives. Non-deterrent enforcement, which is more common in co-management arrangements, may also lead to shifts in strategy by signaling cooperative norms—group-shared standards of behavior that dictate how to act in specific situations (Fehr and Schurtenberger 2018). In such cases, cooperation can become an individually beneficial choice, even for those who do not typically adhere to the cooperative norm, provided they have high expectations that others will comply with it.

All this evidence suggests the distribution of cooperative strategies within a group of CPR users determines how well they can sustain cooperation and consequently could explain variability in co-management outcomes. In addition, it indicates that understanding how external enforcement affects the distribution of cooperative strategies could shed light on crowding-in and out effects and the role of norms and expectations in shaping this interaction. Here we address these two proximate mechanisms (the enforcement of social norms through external enforcement and the role of strategic heterogeneity) as potential explanation for the varying levels and dynamics of cooperation in the co-management of CPRs. To explore these ideas, we use a lab-in-the-field CPR experiment performed with 85 small-scale fishers exploiting benthic resources along the central coast of Chile. The sample includes three types of CPR user groups with differing co-management experiences: fishers from unions participating in high-performing co-management (HP), fishers from unions participating in low-performing co-management (LP), and a group of non-unionized fishers who do not participate in co-management (NU). Published results with this subject pool show that group-level behavioral differences across these groups in the CPR game reflect real-life differences in co-management, accounting for the external validity of the experiment (Gelcich et al. 2013). Here, we assess how strategic heterogeneity across user groups relates to differences in their co-management outcomes and explain cooperation dynamics in the absence of external enforcement. We also examine how the introduction of external enforcement affects cooperation dynamics by shifting the distribution of cooperative strategies within user groups and discuss how differences in responses to enforcement across groups can be explained by their real-life co-management experiences. Analyzing the interplay between individual strategies, external enforcement, and co-management outcomes in an externally valid experiment, provides us with a detailed account of how resource users behavior is influenced by group-developed cooperative mechanisms under co-management institutions. These insights are crucial to understand variability in co-management outcomes in real-world settings.

3.2. RESEARCH SETTING AND SUBJECT POOL

In Chile, fishers who are part of a union can collectively apply for Territorial Use Rights for Fisheries within a defined geographic area referred to as Management and Exploitation Areas for Benthic Resources (MEABRs). In this co-management system, fishing unions are required to collaborate with scientific consultants to develop a management plan that combines conservation and exploitation criteria (Gelcich et al. 2010). If approved by state agencies, the MEABR will come into effect and the fishing union will have the exclusive right to exploit benthic resources in the specified area and bear the responsibility of monitoring compliance with the management plan, which will be regularly enforced by the National Fisheries and Aquaculture Service (SERNAPESCA). On the other hand, non-unionized fishers can exploit benthic resources outside these protected areas, but they can engage in illegal poaching within MEABRs, causing harm to both the ecological and economic yields of fishing unions (Oyanedel et al. 2018, Romero et al. 2022).

Since 1997, more than seven hundred MEABRs have been designated. Regarding their effectiveness in enhancing local livelihoods, scientific assessments have thus far emphasized that the success of MEABRs depends on environmental dynamics (Aburto et al. 2013, Anguita et al. 2020), and users' social and risk preferences (Gelcich et al. 2007). Furthermore, user perceptions are intertwined with the outcomes of co-management, encompassing expectations of organizational improvements and enhanced resource management, as well as concerns about increased illegal fishing and economic performance falling short of expectations (González et al. 2006, Gelcich et al. 2007, Aburto et al. 2013, Romero et al. 2022). Although there is evidence that, overall, belonging to a union and engaging in the collective action in MEABRs increase the income of fishers, there is also evidence of significant variance on the effect of MEABRs due to differences in the capacity to control illegal fishing (Romero and Melo 2021).

Another feature that conditions the success of MEABRs is the social-ecological setting. The great extension of the country's coastline (which extends over 6,000 kilometers and crosses several types of ecosystems, resources, and technologies) is inhabited by human populations with different social and cultural backgrounds. Related evidence highlights the existence of differences between large biogeographic regions along the coast of Chile, each exhibiting distinctive characteristics. Additionally, there is heterogeneity within these large regions among neighboring fishing coves in terms of fishing effort, biomass, composition of landings, and per capita income. This intra-regional heterogeneity is likely due to socioeconomic factors such as price fluctuations, individual fishing behaviors, and historical factors (Chevallier et al. 2021). To control for the large biological and sociocultural variability, we selected a subject pool from a limited area of the central coast of Chile. This ensured a culturally similar population, sharing the same geography and ecosystem, but with organizational differences. All experimental subjects live in semi-urban settlements of the Valparaíso and O'Higgins administrative regions in the central coast of Chile. Small-scale fisheries in these regions are primarily oriented toward commerce rather than direct consumption and are characterized by a blend of traditional and modern techniques and technologies. Commercially traded marine species include hake (*Merluccius gayi gayi*), golden kinglip (*Genypterus blacodes*), Jack mackerel (*Trachurus murphyi*) mollusks such as Loco (*Concholepas concholepas*) and Keyhole limpets (*Fissurella spp*), and sea urchins (*Loxechinus albus*). Fishing activities are also complemented by gastronomy, tourism, and occasionally farming. Unlike the northern and southern regions of Chile, this section of the central coast exhibits a low engagement of indigenous population in official fisheries management.

The CPR experiment was performed with 85 small-scale fishers, with 30 belonging to three fishing unions that exhibit high performance in MEABR management (HP, n = 30), 25 belonging to two fishing unions displaying relatively low performance in the management of their MEABRs (LP, n = 25), and 30 non-unionized fishers who do not participate in MEABR management (NU, n = 30). Classification of unions into high- or low-performance was based on data collected by Marín et al. (2012) and considered variables such as compliance with norms and rules of each union, co-management performance measures (assessed by the National Fisheries and Aquaculture Service), and ecological outcomes (measured as the evolution of Total Allowable Catch and from biodiversity levels: see appendix 1 for details on this classification).

High-performance unions (HP) are efficient organizations with stable levels of cooperation and sustainable exploitation of benthic resources. They feature livelihood diversification (i.e., having a wide range of targeted resources, knowledge, and technologies for complementary sources of income, besides their MEABR), and clear and valued mechanisms for the creation and enforcement of norms discussed in monthly meetings, such as harvesting plans and graduated sanctions for norm transgressions. These organizational-level features are associated with profitable and sustainable ecological outcomes (e.g., sustained or increased total allowable catches over the last five years). Low-performance unions (LP) lack HP attributes and exhibit deficient social and environmental results. LP unions do not have formal MEABR committees that meet monthly, and neither possess formal or informal mechanisms for norm enforcement. Moreover, their members developed a narrative justifying overextraction. These characteristics reflect low esteem on unions and MEABRs, and resulted in unprofitable and unsustainable MEABRs (e.g., with decreasing total allowable catch over the last five years). Finally, non-unionized fishers were recruited from nearby localities of HP and LP unions. While they understand co-management advantages, they decided not to participate in unions or the MEABR system. They extract benthic resources on their own in open-access areas and frequently poach into unions' MEABR. The reasons for not participating in the MEABR system include narratives against this institution as usurping historical open-access zones and unwillingness to bear the costs of collective action.

Fishers in Central Chile base their activities in coves near their settlements, where they land their catch, store their gear, and socialize. It is common for unions to be associated with one of these coves which they may share with other unions or non-unionized fishers. All unions in our sample come from different coves although it is possible that members of different unions live in the same locality as well as non-unionized fishers. In Table 1 we present demographic and geographic characteristics of the six unions sampled for this study. In general terms, HP and LP unions are similar in the number of years working on their MEABR (mean 11.6 years for HP and mean 8.3 years for LP); MEABRs surface (mean 111 ha for HP and mean 131 ha for LP); nearest city from the centroid of MEABRs (mean 59 km for HP and mean 77.6 km for LP); and nearest cove from the centroid of MEABRs (mean 0.6 km for HP and mean 0.7 km for LP). There are larger differences in the number of members (mean 53.5 fishers for HP and mean 32.3 fishers for LP), and in the union's age (mean 33 years for HP and mean 17 years for LP). We attempted to balance the sample so that each type of user group (HP and LP) included unions with varying levels of dependence on benthic resources, while ensuring that the remaining characteristics were as homogeneous as possible.

Table 1. HP and LP user groups demographics.

Union	Performance	Union's age at fieldwork's date	Years of MEABR until the date of fieldwork	Number of members	MEABR surface (ha)	Nearest city (km)	Nearest cove (km)	Dependency on benthic resources
A	HP	23	13	40	108.18	43.1	0.05	High (75%)
B	HP	53	13	93	204.14	33.9	0.33	Medium (23%)
C	HP	24	9	27	21.34	100.9	1.52	Low (3%)
D	LP	8	8	39	177.93	99.2	0.17	Low (3%)
E	LP	21	8	19	193.45	41.2	0.73	High (74%)
F	LP	23	9	39	22.65	92.4	1.26	Medium (25%)

3.3. EXPERIMENTAL PROCEDURE

The CPR experiment had a within-subjects design in which all 85 individuals were exposed to two treatments: an unenforced treatment (rounds 1 to 10) and an enforced treatment (rounds 11 to 20). In the experiment, subjects were assembled randomly into fixed groups of five individuals from the same fishing union. In the case of non-unionized fishers, they played with other non-unionized fishers from nearby localities. The experiment implemented anonymous protocols to reduce third-party observation effects and reputational concerns. Additionally, no communication was allowed, payments were made privately at the end of each session, and the instructions were contextually framed to resemble their real-life fishing activities (Appendix 2 to see game instructions).

Before starting the experimental procedure, subjects answered a questionnaire and played five test rounds to confirm task comprehension. After this, the real experiment began according to the following procedure: each subject $i \in \{1,\ldots,5\}$ was informed that she possessed, at the beginning of each round, an endowment of 100 units of a local benthic fishery resource known as 'Loco' (*Concholepas concholepas*). We framed these 100 units as their individual quota. Then, in every round, each subject privately decided whether she wanted to harvest above their quota up to 50 additional units of Loco. The overharvest is then defined as $y_{it} \in \{0,..., 50\}$, and measures the extra units of Loco extracted by i in round t. Each over-harvested unit increases individual payments but also implies that each other member of the group loses half a 'Loco' (resembling the real collective negative externalities associated with the overexploitation and depletion of resource stocks). During the first ten rounds of the game (the unenforced treatment) there were no sanctions to $y_{it} > 0$, and players only had to choose their over-extraction level per round. After each round, subjects were informed of i) the average harvest of their group, ii) their individual payoff of that round, and iii) their individual losses due to the group's over-harvested Locos. No information was given about teammates' individual over extracted levels and, therefore, individual behavior was private for every player in the experiment. Each 'Loco' earned was worth $10 CLP (10 Chilean pesos), resulting in individual payoffs between rounds 1 to 10 given by:

Equation 1

$$\pi_{it} = \$10\left(100 + y_{it} - \frac{1}{2}\sum_{j-i} x_{jt}\right)$$

Where π_{it} is the payoff of subject *i* in round *t*, y_{it} is the over extraction of subject *i* at round *t*, and x_{jt} is the sum of units over-extracted by *i*'s teammates at round *t*. In this setup, and considering only self-interested rational individuals, the only equilibrium of the game is a Tragedy of the Commons where every player over-extract 50 units per round, resulting in a Pareto-inefficient outcome.

The rules changed for rounds 11 to 20 (the enforced treatment) with the introduction of an external non-deterrent enforcement mechanism, implemented as an economic punishment. Under this new setup, two subjects per group were randomly selected after each round for inspection, resembling an external enforcement mechanism implemented by a state agency. If the inspected subject exceeded her individual quota (i.e., if $y_{it} > 0$), all her extraction in that round was "confiscated" and the subject lost the entire round payoff. After each round, subjects were individually informed of the average extraction of the group, their payoff loss due to the group's over-extraction, and their payoff. Therefore, for rounds 11 to 20, the payoff function for subject *i* can be formalized as:

Equation 2

$$\begin{cases} \pi_{it} = \$10(100 + y_{it} - \frac{1}{2}\sum_{j-i} x_{jt}), if\ y_{it} = 0 \\ \pi_{it} = \frac{3}{5}\$10(100 + y_{it} - \frac{1}{2}\sum_{j-i} x_{jt}), if\ y_{it} > 0 \end{cases}$$

This enforcement mechanism alone is not sufficient to deter overextraction. The incentive was designed so the decision to comply with the quota depended on the subject's expectations on others' overextraction. For a subject to dismiss over-extracting in round *t*, it must expect:

Equation 3

$$E\sum_{j-i} x_{jt} \leq 50$$

where the left side of the inequation expresses the expectation that subject *i* has on their teammates' overextraction in round *t*. Therefore, a subject will dismiss over extracting only if she expects that her teammates will over-extract 50 or fewer units of loco in round *t*; if this condition is not met, she might find no incentives to restrain herself from over-extraction.

Our design combines ecological validity (using repeated interactions, an external monetary enforcement mechanism, and framed instructions, to resemble the conditions that subjects face daily) and a sample with external validity (through the prior classification of the subject pool based on social and biological performance), favoring a better understanding of the phenomenon under study.

3.4. STATISTICAL ANALYSIS

To identify the strategy of each of the 85 experimental subjects in each treatment, we employed two procedures, similar to those already performed by Kurzban and Houser (2005) and van Klingeren (2022). First, we identified the

Linear Conditional Contribution Profile (LCCP) of each subject in both treatments using an econometric model. Then, we used the LCCP results to classify each subject into a cooperative type in each treatment. The LCCP captures the levels of unconditional and conditional cooperation for each subject by assessing how they adjust their overharvest decisions in response to others' behavior over the rounds of the game. We define unconditional cooperation as the level of cooperation subjects display when there is no cooperation from others, and conditional cooperation as the extent to which subjects adjust their cooperation to match the cooperative behavior of their peers.

To interpret decisions in terms of cooperation, we transformed the overextraction levels of each subject into an individual normalized cooperation per round, i.e., (50 - y_{it}) / 50, resulting in compliance per round ranging from 0 to 1 (that is, an overharvest of 50 units in round *t* is equal to a normalized compliance of 0 for that round, and no overextraction corresponds to a normalized compliance of 1). Next, we estimated the LCCP of each subject according to:

Equation 4

$$y_{it} = \alpha_{i1} + \alpha_{i2}\mathrm{DR} + \beta_{i1}\bar{x}_{-it-1} + \beta_{i2}\bar{x}_{-it-1}\mathrm{DR} + \varepsilon_{it}$$

Where y_{it} is the compliance of subject *i* in round *t*, DR is a dummy variable that takes the unenforced treatment as a reference category, $\bar{\mathrm{x}}_{-it-1}$ is the average compliance of *i*'s teammates in *t-1*, and ε_{it} is an error term. Rounds 1 and 11 were dropped from regressions since there is no *t-1* in round 1 and *t-1* in round 11 has a trend of the unenforced treatment and could have distorted our results. The estimated parameters of our model can be interpreted as:

$\widehat{\boldsymbol{\alpha}}_{i1}$ is the unconditional contribution of subject *i* in the unenforced treatment

$\widehat{\boldsymbol{\alpha}}_{i2}$ is the change in the unconditional contribution of subject *i* due to the introduction of the enforced treatment

$\widehat{\boldsymbol{\alpha}}_{i1} + \widehat{\boldsymbol{\alpha}}_{i2}$ is the unconditional contribution of subject *i* in the enforced treatment

$\widehat{\boldsymbol{\beta}}_{i1}$ measures the responsiveness of subject *i* to *-i's* compliance in the unenforced treatment

$\widehat{\boldsymbol{\beta}}_{i2}$ is the change in the responsiveness of subject *i* due to the introduction of the enforced treatment

$\widehat{\boldsymbol{\beta}}_{i1}$+ $\widehat{\boldsymbol{\beta}}_{i2}$ is the responsiveness of subject *i* to *-i's* compliance in the enforced treatment

Note that similar models have been previously employed in studies exploring strategic heterogeneity. These models employ linear regressions with individual contributions regressed against contributions from other team members, and their parameters bear a similar interpretation to ours, where the intercept represents the level of unconditional cooperation, and the slope represents the level of conditional cooperation (Carpenter and Seki, 2011; Kurzban and Houser, 2005; van Klingeren, 2022).

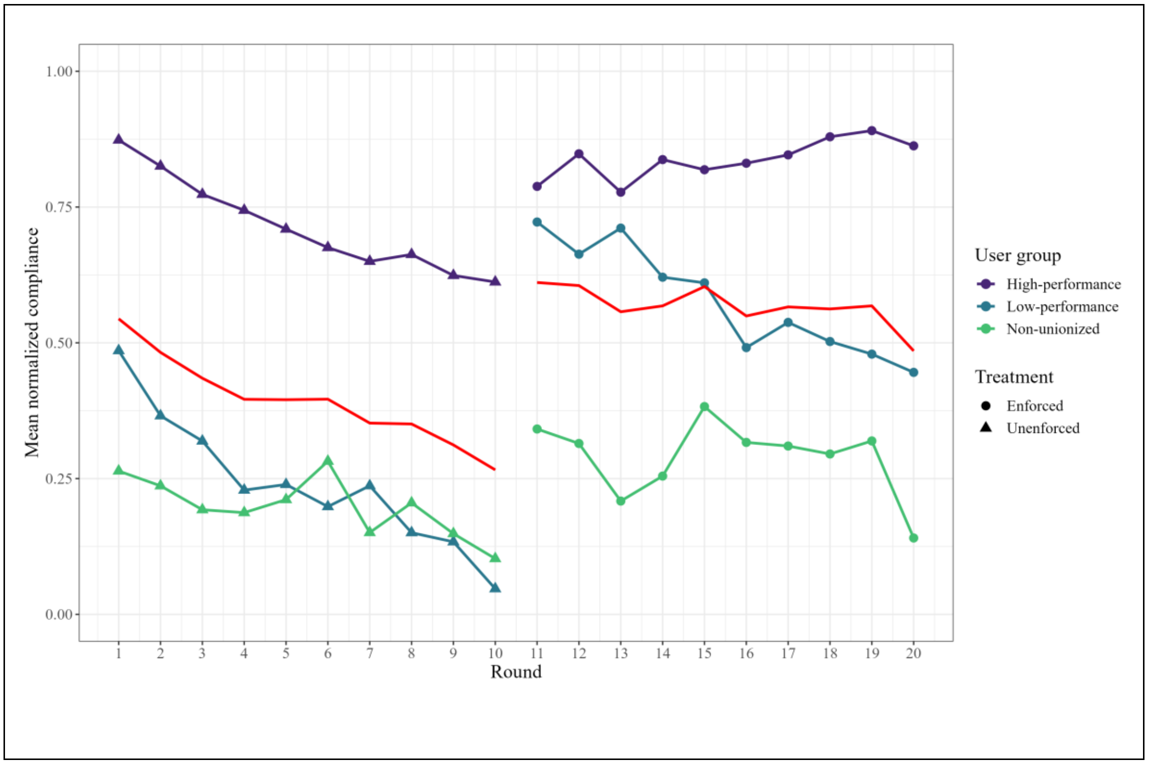
We estimated these parameters using Constrained Least Squares, which imposes restrictions to the minimization of the sum of squares residual in a linear model. In our case, we defined constraints as lower bounds of 0 and upper bounds of 1 to a set of parameters, to ensure that the LCCPs were always contained in the feasible region of the model (where compliance values less than 0 or greater than 1 were not possible). These bound were finally defined for: $\widehat{\alpha}_{i1}$; $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2}$; $\widehat{\alpha}_{i1} + \widehat{\beta}_{i1}$; and $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2} + \widehat{\beta}_{i1} + \widehat{\beta}_{i2}$.

We then used the LCCP results to classify each subject into free riders, conditional cooperators, unconditional cooperators, or negative cooperators by adapting the methods of Kurzban and Houser (2005) and van Klingeren (2022). We classified each subject in each treatment to examine the distribution of strategies among different user groups (high-performers, low-performers and non-unionized), to understand how external enforcement influenced individual strategies, and to explore the impact of these strategy changes on the erosion or stabilization of cooperation over successive rounds. In our classification scheme, a free-rider is a subject whose LCCP is always lower than half of the total possible compliance (i.e., their expected compliance is always < 0.5 at all levels of others' contributions); on the contrary, subjects classified as unconditional cooperators always display an LCCP above half of the total possible compliance (i.e., their expected compliance is always > 0.5 at all levels of others' contributions); conditional cooperators comply below 0.5 when their teammates' compliance is absent (i.e, the intercept of their LCCP is < 0.5), but the responsiveness to their teammates is above 0.5 (intercept plus slope > 0.5); and negative cooperators comply above 0.5 when there is no compliance from others (i.e., their LCCP intercept is > 0.5), but the responsiveness to their teammates is below 0.5 (intercept plus slope < 0.5. See Appendix 3 to a more detailed and formal description of the classification scheme).

3.5. RESULTS

The dynamics of average compliance across rounds accounts for the externally valid result reported in Gelcich et al. (2013). Figure 2 shows that in the unenforced treatment (rounds 1 to 10) compliance declined over rounds in all three user groups, while the introduction of the external enforcement (rounds 11 to 20) generated different effects: i) HP unions recovered their initial levels of compliance (comparing rounds 11 and 1), and, more importantly, were the only ones capable of stabilizing cooperation until the end of the game; ii) LP unions increased their initial compliance levels (comparing rounds 11 and 1), but were unable to stabilize it; and iii) non-unionized fishers slightly increased their initial compliance, but were also incapable of sustaining it until the end of the game. To assess the compliance differences between the initial rounds of both treatments (rounds 1 and 11), we conducted a Wilcoxon signed-rank test that suggests statistically significant differences only for the HP unions ($V=27.5$, $p=0.038$), while no such significance is observed for the LP unions($V=79.5$, $p=0.279$) or non-unionized fishers ($V=53.5$, $p=0.17$).

Figure 1. Normalized mean compliance of user groups across rounds. Mean compliance across user groups in red.



This result not only supports the external validity of the framed CPR experiment in our subject pool, but also enables us to establish HP unions in the enforced treatment as a benchmark for user groups capable of stabilizing cooperation.

Due to problems in the variance of the independent variables, we were forced to drop seven individuals (five from HP and two LP unions); thus, the final sample was fixed at HP, n = 23, LP, n = 24, and NU, n=30. The remaining 77 LCCPs were properly estimated and are presented, case by case, in the supplementary material (Appendix 4). In Tables 2 to 4, we present transition matrices that show changes in the distribution of strategic types between the unenforced and enforced treatment for each user group (Appendix 5 to the whole sample strategy distribution). We conducted goodness-of-fit Chi-squared tests to test the equality of strategy proportions within each user group-treatment combination. All the tests yielded statistically significant outcomes, indicating that for each user group-treatment combination, the strategy proportions are significantly distinct.

For HP unions, the distribution of strategies under the unenforced treatment shows an almost equal proportion of conditional and unconditional cooperators (Table 2). This leads to a high initial compliance, followed by a decline over the course of the treatment rounds. The implementation of external enforcement significantly shifts strategy proportions within HP unions, with nearly 70% of subjects adopting an unconditional cooperation strategy, less than 20% engaging in conditional cooperation, and less than 10% representing free-riders and negative cooperators. The distribution of HP unions in the enforced treatment exhibits great asymmetry among strategy proportions, dominated by unconditional cooperators. For LP unions, the distribution in the unenforced treatment shows a nearly equivalent proportion between conditional cooperators and free-riders, with a marginal incidence of the remaining strategies (Table 3).

Consequently, within this treatment, there is a dominance of merely two strategies, which in combination constitute over 90% of the proportion. The impact of external enforcement leads to a significant change, primarily by halving the percentage of free-riders and augmenting the prevalence of unconditional cooperators. In this treatment, the distribution of proportions appears much less even than under the unenforced treatment. Lastly, for non-unionized fishers, the distribution of strategies under the unenforced treatment exhibits two distinct characteristics: the prevalence of free-riders (with over 75% of subjects displaying this behavior) and the absence of unconditional cooperators (Table 4). This distribution is among the most unequal and its composition explains the low level of compliance between rounds 1 and 10. External enforcement impacts these proportions, reducing the incidence of free-riders by almost 20% and increasing negative cooperators by around 15%. The highest proportion of negative cooperators across user groups and treatments is observed for non-unionized fishers under external enforcement, reaching more than 70% of the strategy composition when coupled with free-riders.

Table 2. Effect of external enforcement on the distribution of strategic types in High-Performance unions.

Strategies	CC	FR	NC	UC	Unenforced distribution
CC	8.70% (2)	0.00% (0)	4.35% (1)	26.09% (6)	39.13% (9)
FR	4.35% (1)	4.35% (1)	0.00% (0)	0.00% (0)	8.70% (2)
NC	4.35% (1)	4.35% (1)	0.00% (0)	0.00% (0)	8.70% (2)
UC	0.00% (0)	0.00% (0)	0.00% (0)	43.48% (10)	43.48% (10)
Enforced distribution	17.39% (4)	8.70% (2)	4.35% (1)	69.57% (16)	100.00% (23)

** CC: Conditional cooperators; FR: Free-riders; NC: Negative cooperators; UC: Unconditional cooperators. Column "Unenforced distribution" gives the proportion of each strategy in this type of user group for the unenforced treatment (rounds 1-10); "Enforced distribution" row gives the strategy distribution in the enforced treatment (rounds 11-20). Number of cases inside parenthesis.

Table 3. Effect of external enforcement on the distribution of strategic types in Low-Performance unions.

Strategies	CC	FR	NC	UC	Unenforced distribution
CC	16.67% (4)	8.33% (2)	4.17% (1)	12.50% (3)	41.67% (10)
FR	25.00% (6)	16.67% (4)	0.00% (0)	8.33% (2)	50.00% (12)
NC	4.17% (1)	0.00% (0)	0.00% (0)	0.00% (0)	4.17% (1)
UC	0.00% (0)	0.00% (0)	0.00% (0)	4.17% (1)	4.17% (1)
Enforced distribution	45.83% (11)	25.00% (6)	4.17% (1)	25.00% (6)	100.00% (24)

Table 4. Effect of external enforcement on the distribution of strategic types in non-unionized fishers.

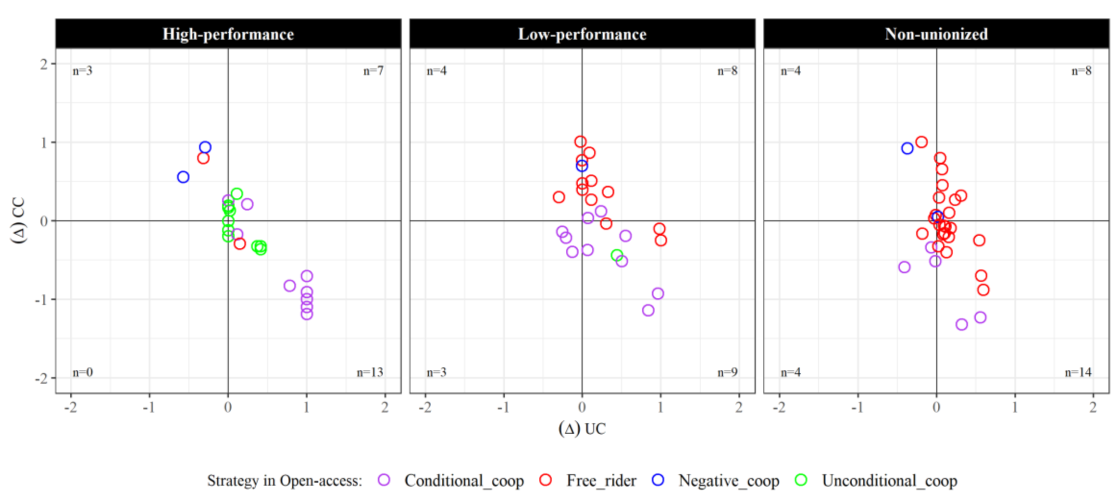
Strategies	CC	FR	NC	UC	Unenforced distribution
CC	3.33% (1)	10.00% (3)	3.33% (1)	0.00% (0)	16.67% (5)
FR	16.67% (5)	43.33% (13)	13.33% (4)	3.33% (1)	76.67% (23)
NC	3.33% (1)	0.00% (0)	3.33% (1)	0.00% (0)	6.67% (2)
UC	0.00% (0)	0.00% (0)	0.00% (0)	0.00% (0)	0.00% (0)
Enforced distribution	23.33% (7)	53.33% (16)	20.00% (6)	3.33% (1)	100.00% (30)

Table 5. Goodness-of-fit Chi-squared tests by CPR user groups and treatment.

CPR user groups	Unenforced treatment	Enforced treatment
High-performance	$X^2(3, N = 23) = 9.86, p = 0.019$	$X^2(3, N = 23) = 25.17, p = 1.42e-05$
Low-performance	$X^2(3, N = 24) = 17, p = 0.000$	$X^2(3, N = 24) = 8.33, p = 0.039$
Non-unionized	$X^2(2, N = 30) = 25.8, p = 2.498e-06$	$X^2(3, N = 30) = 15.6, p = 0.001$

Levels of conditional and unconditional cooperation for each user group are illustrated in Figure 2. In this plot, changes in each level are the coefficients $\widehat{\alpha}_{i2}$ (ΔUC) and $\hat{\beta}_{i2}$ (ΔCC) of each subject's LCCP. This result shows that, within HP unions, individuals classified as unconditional cooperators in the unenforced treatment hardly alter their strategies once enforcement is introduced. On the other hand, many conditional cooperators increase their levels of unconditional cooperation (on the right side in the x-axis) and decrease their conditionality (down along the y-axis). This small cluster of subjects is the group that succeeds in altering the average behavior of HP unions in the enforced treatment and manages to stabilize compliance across rounds. In the case of LP unions, there is a greater tendency for free-riders to increase their levels of conditionality (moving up along the y-axis) as a response to the enforced treatment, while conditional cooperators show a decrease in their levels of conditionality (moving down along the same y-axis). On the other hand, NU fishers have several subjects clustered near the origin of the plot, demonstrating the limited impact of enforcement on these individuals.

Figure 2. Change of unconditional and conditional cooperation levels by CPR user groups due to enforcement. Colors indicate classification in the unenforced treatment.



3.6. DISCUSSION

In this study, we analyze the results of an ecologically valid common pool resource experiment involving artisanal fishers who exploit benthic resources along the coast of Chile, aiming to determine whether cooperative strategies and external enforcement could account for differences in co-management outcomes across user groups. Our findings reveal that the cooperation dynamics observed in the experiment, which reflect co-management outcomes of different user groups, can be attributed to the distribution of cooperative strategies and their interaction with external enforcement. When viewed through the lens of experimental economics, differences in the strategic composition of user groups and their responsiveness to external enforcement are likely shaped by variations on how groups have developed cooperative norms within co-management. Our results shed light on how norm enforcement affects individuals' strategy-switching behavior, which is crucial for understanding the dynamics of cooperation in commons problems, such as those faced by fishers in co-management regimes.

Aggregated experimental findings reveal cooperation dynamics in the game that reflect differences in users' group ability to sustain cooperation in real-life co-management. Specifically: a) unions engaged in high-performing co-management experiences displayed elevated cooperation at the beginning of an unenforced CPR game, which subsequently attenuates to moderate levels. However, upon the introduction of enforcement, cooperation reverts to its initial level and remains consistently high; b) unions engaged in low-performing co-management start the unenforced CPR game with moderate cooperation, which erodes over repeated interactions. In the enforced treatment, their initial cooperation mirrors that of high-performing unions in round 11 but erodes over time; c) non-unionized fishers, who have no co-management experience, exhibit minimal cooperation in the unenforced treatment, which marginally rises after the institutional transition.

Considering these cooperation dynamics, four fundamental questions arise. The response to each of these questions is based on the strategic distribution of fishers' behaviors induced by the enforced treatment.

I. **What causes the diminishing cooperation in both high and low-performing unions under the unenforced treatment**? In this treatment, both HP and LP unions predominantly consist of conditional cooperators, which makes cooperation unstable. The difference in initial cooperation levels between HP and LP unions is attributable to the higher proportion of free-riders in low-performing unions.

II. **What causes the resurgence of cooperation to peak levels in both union types when unenforced changes to enforced?** The answer to this question is related to the norm signaling caused by the introduction of external enforcement (McAdams 2000, Fehr and Schurtenberger 2018). This allows HP unions to revert to initial cooperation levels and LP unions to increase the percentage of cooperators (from around 50% to over 70%), due to a reduction in the number of free-riders and an increase in the number of unconditional and conditional cooperators.

III. **Why does the enforced treatment not affect average cooperation in non-unionized fishers?** In this case, the unconditional and conditional cooperators percentage for non-unionized fishers stabilizes below 25%, preventing a cooperative equilibrium.

IV. **Why does cooperation only stabilize in groups from high-performing unions?** These results are caused by the redistribution of strategic types within the different user groups. LP unions witnessed a decline in free-riders by half. However, most of these free-riders transitioned to conditional cooperators, increasing the level of conditional cooperators from ~40% to ~50%. In stark contrast, HP union not only increased the number of cooperators to almost 90%, but also underwent a change in the predominant strategy. Under the enforced treatment, conditional and unconditional cooperators, previously present in similar proportions now constitute around 80% of all cooperators. This suggests a widespread internalization of norms, thereby stabilizing cooperation at elevated levels.

Previous literature on strategic heterogeneity has suggested that greater shares of conditional cooperators in game experiments (compared to free-riders) is associated with greater extents of overall cooperation or productivity in real-life CPRs (Rustagi et al. 2010, Carpenter and Seki 2011). We propose a more nuanced interpretation of the association between the percentage of conditional cooperation and successful CPR management or, more generally, of sustaining cooperation in social dilemmas. In our results, only fishers belonging HP unions in the enforced treatment were capable of stabilizing compliance, while fishers from LP unions and nonunionized fishers failed to do so. HP unions in the unenforced treatment and LP unions in the enforced treatment are characterized by a larger share of conditional cooperators and faced the decline of cooperation over time.

External enforcement played a crucial role in shaping the strategic composition that led to cooperation dynamics in the game that mirror real-life co-management outcomes. Our individual level regressions show that external enforcement mostly led to switches from conditional cooperators to unconditional cooperation strategies in HP unions, similarly to what was found by Rodriguez-Sickert and colleagues (2008). In our sample, the differential effects of external non-deterrent enforcement on strategic composition across groups can be explained by differences in the internalization of norms and expectations across user groups with varying institutional experiences. HP unions are the most responsive to enforcement, resulting in a higher percentage of unconditional cooperators compared to other strategies and even reducing the proportion of free-riders to less than 10%. We can interpret the responsiveness of HP user groups through a normative effect. The introduction of external enforcement signals a standard of behavior, and for subjects that have internalized cooperation in co-management no matter what others are doing, this can lead them to switch into unconditional cooperators (Fehr and Schurtenberger 2018). Subjects that remain conditional cooperators under external enforcement, will also increase their cooperation if they anticipate a higher share of unconditional cooperators under norm enforcement, stabilizing cooperation. Through their daily interaction groups of users develop norms, expectations and other behavioral heuristics that they apply in similar contexts (Ostrom 1990, Rivera-Hechem et al. 2021). Real-life interactions of HP unions under functioning co-management that often involves sanctioning and social pressure, has likely led to norm internalization and high cooperation expectations in contexts that resemble co-management interactions.

While enforcement led to switches in strategies in LP unions increasing unconditional cooperators and decreasing the proportion of free-riders, it fails to sustain the levels of compliance between rounds 11 and 20. This erosion of cooperation in the enforced treatment (which sets it apart from the high-performers) can be attributed to the repeated interaction between these highly represented strategies within LP unions (46% of conditional cooperators and 25% of free-riders. As the low cooperation from free-riders undermines enforcement, resulting in the reduction or withdrawal of compliance from those employing the conditional strategy (de Oliveira et al. 2015). A necessary condition for some conditional cooperators to shift to unconditional cooperation is the belief that external enforcement is sufficient, as the fines imposed on defectors would relieve them of the need to retaliate by withholding their own cooperation (Andreoni

1995). One plausible interpretation for the persistence of a subset of individuals as conditional cooperators is that they maintained this strategy because they perceived the enforcement mechanism as inadequate or ineffective in punishing free-riders, reflecting a potential lack of confidence in external regulators in real-life contexts. In LP communities, these expectations are likely shaped by their real-world experiences of interacting within unions that have low peer accountability, which leads to weak norm internalization.

The situation of non-unionized fishers also embodies individuals who have conducted their activities independently, lacking collective fishing experience or familiarity with institutions such as MEABRs or quota compliance. For this reason, their behavior in the enforced treatment reflects the potential response of a human group to their first experience of collective action regulation, and, in this regard, their low compliance and declining behavior over time make sense if they lack confidence in this type of regulation. Differences in how groups develop norms and expectations around natural resource use are influenced by the costs and benefits each group faces. Although the unions in our sample share similar ecological and cultural backgrounds, other factors known to affect collective action in common-pool resource use may explain why developing cooperation mechanisms in co-management is more cost-effective for some groups than for others. These factors include union-level variables such leadership, social capital and cohesion, and local knowledge (Ostrom 2009, Gutiérrez et al. 2011, Gelcich et al. 2019, Cinner et al. 2019).

Having controlled for ecological and cultural variability (i.e., by working with a sample within one of the major areas identified by Romero and Melo), we can attribute the differences between types of user groups to the informal institutions that have emerged within each union to explain the variation in individual behaviors. The path dependence of each community and the historical context of these groups are reflected in aspects that have molded their unions, particularly through local informal norms. The interaction of these norms with the regulatory framework prompts varying propensities for individuals to responses to the MEABR regulation. For example, in HP user groups, this manifests as greater confidence in the effects of enforcement on other group members, which transforms individual strategies into those of highly cooperative actors by increasing trust and expectations regarding the prosocial commitment of the entire group. The findings of this study can provide insights into the processes that influence the prospect of success or failure of co-management institutions. We observed that in this sample of fishers, the heterogeneity of strategies plays a significant role in the probabilities of success and the stability of cooperation, a conclusion that aligns with previous experimental research.

In this paper, we focus on understanding the dynamics of cooperation—whether cooperation erodes or stabilizes—as an emergent phenomenon resulting from the interaction of individual strategies within groups. By identifying these underlying mechanisms, we aim to provide insights into how different user groups sustain cooperation (or not) in their real-life CPR experiences. This helps to comprehend the social challenges that user groups face in maintaining cooperation, and propose potential solutions, thereby shedding light on ways to support local communities in engaging more effectively in co-management. Furthermore, results show external enforcement drive strategic shifts in individuals who are more inclined to trust its impact on others' behavior. Future research on the role of heterogeneous strategy distribution and its interaction with the institutional environment to understand the dynamics of cooperation can delve more deeply into the effects of various strategy configurations on group outcomes, as well as explore the marginal effects of increasing proportions of each strategy.

3.7. LITERATURE CITED

Aburto, J., G. Gallardo, W. Stotz, C. Cerda, C. Mondaca-Schachermayer, and K. Vera. 2013. Territorial user rights for artisanal fisheries in Chile – intended and unintended outcomes. *Ocean & Coastal Management* 71:284–295.

Andreoni, J. 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? *The American Economic Review* 85(4):891–904.

Anguita, C., S. Gelcich, M. Aldana, and J. Pulgar. 2020. Exploring the influence of upwelling on the total allowed catch and harvests of a benthic gastropod managed under a territorial user rights for fisheries regime along the Chilean coast. *Ocean & Coastal Management* 195:105256.

Arias Schreiber, M. 2012. The evolution of legal instruments and the sustainability of the Peruvian anchovy fishery. *Marine Policy* 36(1):78–89.

d'Armengol, L., M. Prieto Castillo, I. Ruiz-Mallén, and E. Corbera. 2018. A systematic review of co-managed small-scale fisheries: Social diversity and adaptive management improve outcomes. *Global Environmental Change* 52:212–225.

Baggio, J., A. Barnett, I. Perez-Ibarra, U. Brady, E. Ratajczyk, N. Rollins, C. Rubiños, H. Shin, D. Yu, R. Aggarwal, J. Anderies, and M. Janssen. 2016. Explaining success and failure in the commons: the configural nature of Ostrom's institutional design principles. *International Journal of the Commons* 10(2):417–439.

Berkes, F. 1985. Fishermen and 'The Tragedy of the Commons.' *Environmental Conservation* 12(3):199–206.

Berkes, F., P. George, and R. J. Preston. 1991. Co-management: The Evolution in Theory and Practice of the Joint Administration of Living Resources. *Alternatives* 18(2):12–18.

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

Brown, G., and C. Harris Jr. 2008. National forest management and the "tragedy of the commons": A multidisciplinary perspective. *Society & Natural Resources*.

Burlando, R. M., and F. Guala. 2005. Heterogeneous Agents in Public Goods Experiments. *Experimental Economics* 8(1):35–54.

Cardenas, J. C., J. Stranlund, and C. Willis. 2000. Local Environmental Control and Institutional Crowding-Out. *World Development* 28(10):1719–1733.

Carlsson, L., and F. Berkes. 2005. Co-management: concepts and methodological implications. *Journal of Environmental Management* 75(1):65–76.

Carpenter, J., and E. Seki. 2011. Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay. *Economic Inquiry* 49(2):612–630.

Castilla, J. C., and M. Fernandez. 1998. Small-Scale Benthic Fisheries in Chile: On Co-Management and Sustainable Use of Benthic Invertebrates. *Ecological Applications* 8(sp1):S124–S132.

Chaudhuri, A. 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics* 14(1):47–83.

Cheung, S. L. 2014. New insights into conditional cooperation and punishment from a strategy method experiment. *Experimental Economics* 17(1):129–153.

Chevallier, A., B. R. Broitman, N. Barahona, C. Vicencio-Estay, F. K. C. Hui, P. Inchausti, and W. B. Stotz. 2021. Diversity of small-scale fisheries in Chile: Environmental patterns and biogeography can inform fisheries management. *Environmental Science & Policy* 124:33–44.

Cinner, J. E., J. D. Lau, A. G. Bauman, D. A. Feary, F. A. Januchowski-Hartley, C. A. Rojas, M. L. Barnes, B. J. Bergseth, E. Shum, R. Lahari, J. Ben, and N. A. J. Graham. 2019. Sixteen years of social and ecological dynamics reveal challenges and opportunities for adaptive management in sustaining the commons. *Proceedings of the National Academy of Sciences* 116(52):26474–26483.

Cole, D. H., G. Epstein, and M. D. Mcginnis. 2014. Digging deeper into Hardin's pasture: the complex institutional structure of 'the tragedy of the commons.' *Journal of Institutional Economics* 10(3):353–369.

Cox, M., G. Arnold, and S. Villamayor Tomás. 2010. A Review of Design Principles for Community-based Natural Resource Management. *Ecology and Society* 15(4).

Cronkleton, P., J. M. Pulhin, and S. Saigal. 2012. Co-management in Community Forestry: How the Partial Devolution of Management Rights Creates Challenges for Forest Communities. *Conservation and Society* 10(2):91–102.

Defeo, O., M. Castrejón, R. Pérez-Castañeda, J. C. Castilla, N. L. Gutiérrez, T. E. Essington, and C. Folke. 2016. Co-management in Latin American small-scale shellfisheries: assessment from long-term case studies. *Fish and Fisheries* 17(1):176–192.

Evans, L., N. Cherrett, and D. Pemsl. 2011. Assessing the impact of fisheries co-management interventions in developing countries: A meta-analysis. *Journal of Environmental Management* 92(8):1938–1949.

Fehr, E., and I. Schurtenberger. 2018. Normative foundations of human cooperation. *Nature Human Behaviour* 2(7):458–468.

Fischbacher, U., and S. Gächter. 2010. Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review* 100(1):541–556.

Fischbacher, U., S. Gächter, and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71(3):397–404.

Fleischman, F. D., B. Loken, G. A. Garcia-Lopez, and S. Villamayor-Tomas. 2014. Evaluating the utility of common-pool resource theory for understanding forest governance and outcomes in Indonesia between 1965 and 2012 8(2):304.

Frey, U. J. 2017. Cooperative strategies outside the laboratory — evidence from a long-term large-N-study in five countries. *Evolution and Human Behavior* 38(1):109–116.

Gautam, A. P., and G. P. Shivakoti. 2005. Conditions for Successful Local Collective Action in Forestry: Some Evidence From the Hills of Nepal. *Society & Natural Resources* 18(2):153–171.

Gelcich, S., G. Edwards-Jones, and M. J. Kaiser. 2007. Heterogeneity in fishers' harvesting decisions under a marine territorial user rights policy. *Ecological Economics* 61(2):246–254.

Gelcich, S., R. Guzman, C. Rodríguez-Sickert, J. C. Castilla, and J. C. Cárdenas. 2013. Exploring External Validity of Common Pool Resource

Experiments: Insights from Artisanal Benthic Fisheries in Chile. *Ecology and Society* 18(3).

Gelcich, S., T. P. Hughes, P. Olsson, C. Folke, O. Defeo, M. Fernández, S. Foale, L. H. Gunderson, C. Rodríguez-Sickert, M. Scheffer, R. S. Steneck, and J. C. Castilla. 2010. Navigating transformations in governance of Chilean marine coastal resources. *Proceedings of the National Academy of Sciences* 107(39):16794–16799.

Gelcich, S., M. J. Martínez-Harms, S. Tapia-Lewin, F. Vasquez-Lavin, and C. Ruano-Chamorro. 2019. Comanagement of small-scale fisheries and ecosystem services. *Conservation Letters* 12(2):e12637.

Gneezy, U., and A. Rustichini. 2000. A Fine Is a Price. *The Journal of Legal Studies* 29(1):1–17.

González, J., W. Stotz, J. Garrido, J. M. Orensanz, A. M. Parma, C. Tapia, and A. Zuleta. 2006. The Chilean TURF system: how is it performing in the case of the loco fishery?

Gutiérrez, N. L., R. Hilborn, and O. Defeo. 2011. Leadership, social capital and incentives promote successful fisheries. *Nature* 470(7334):386–389.

Hardin, G. 1968. The Tragedy of the Commons. *Science* 162(3859):1243–1248.

Herrmann, B., and C. Thöni. 2009. Measuring conditional cooperation: a replication study in Russia. *Experimental Economics* 12(1):87–92.

van Klingeren, F. 2022. Using player types to understand cooperative behaviour under economic and sociocultural heterogeneity in common-pool resources: Evidence from lab experiments and agent-based models. *PloS One* 17(5):e0268616.

Klooster, D. 2000. Institutional Choice, Community, and Struggle: A Case Study of Forest Co-Management in Mexico. *World Development* 28(1):1–20.

Kocher, M. G., T. Cherry, S. Kroll, R. J. Netzer, and M. Sutter. 2008. Conditional cooperation on three continents. *Economics Letters* 101(3):175–178.

Kraak, S. B. M. 2011. Exploring the ‘public goods game’ model to overcome the Tragedy of the Commons in fisheries management. *Fish and Fisheries* 12(1):18–33.

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

Ledyard, J. O. 1995. Public Goods: A Survey of Experimental Research. Pages 111–194 *in* J. H. Kagel and A. E. Roth, editors. Princeton University Press, Princeton.

Marín, A., S. Gelcich, J. Castilla, and F. Berkes. 2012. Exploring Social Capital in Chile’s Coastal Benthic Comanagement System Using a Network Approach. *Ecology and Society* 17(1).

McAdams, R. H. 2000. The focal point theory of expressive law. Page *Encyclopedia of Law and Economics*. Edward Elgar Publishing Limited.

Naar, N. 2020. Gaming Anthropology: The Problem of External Validity and the Challenge of Interpreting Experimental Games. *American Anthropologist* 122(4):784–798.

de Oliveira, A. C. M., R. T. A. Croson, and C. Eckel. 2015. One bad apple? Heterogeneity and information in public good provision. *Experimental Economics* 18(1):116–135.

Ostrom, E. 1990, November. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press. https://www.cambridge.org/core/books/governing-the-commons/7AB7AE11BADA84409C34815CC288CD79.

Ostrom, E. 1998. A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997. *American Political Science Review* 92(1):1–22.

Ostrom, E. 2006. The value-added of laboratory experiments for the study of institutions and common-pool resources. *Journal of Economic Behavior & Organization* 61(2):149–163.

Ostrom, E. 2009. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science* 325(5939):419–422.

Oyanedel, R., A. Keim, J. C. Castilla, and S. Gelcich. 2018. Illegal fishing and territorial user rights in Chile. *Conservation Biology* 32(3):619–627.

Pisor, A. C., M. M. Gervais, B. G. Purzycki, and C. T. Ross. 2020. Preferences and constraints: the value of economic games for studying human behaviour. *Royal Society Open Science* 7(6):192090.

Plummer, R., and J. Fitzgibbon. 2004. Co-management of Natural Resources: A Proposed Framework. *Environmental Management* 33(6):876–885.

Podimata, M. V., and P. C. Yannopoulos. 2015. Evolution of Game Theory Application in Irrigation Systems. *Agriculture and Agricultural Science Procedia* 4:271–281.

Pomeroy, R. S., and F. Berkes. 1997. Two to tango: The role of government in fisheries co-management. *Marine Policy* 21(5):465–480.

Rivera-Hechem, M. I., R. A. Guzmán, C. Rodríguez-Sickert, and S. Gelcich. 2021. Effects of experience with access regimes on stewardship behaviors of small-scale fishers. *Conservation Biology* 35(6):1913–1922.

Rivera-Hechem, M. I., C. Rodríguez-Sickert, R. A. Guzmán, T. Ramírez-Parada, F. Benavides, V. Landaeta-Torres, M. Aspé-Sánchez, and G. M. Repetto. 2020. No association between genetic variants in MAOA, OXTR, and AVPR1a and cooperative strategies. *PLOS ONE* 15(12):e0244189.

Rodriguez-Sickert, C., R. A. Guzmán, and J. C. Cárdenas. 2008. Institutions influence preferences: Evidence from a common pool resource experiment. *Journal of Economic Behavior & Organization* 67(1):215–227.

Romero, P., R. A. Estévez, P. Romero, and S. Gelcich. 2022. Artisanal Fisher Association Leaders' Estimates of Poaching in Their Exclusive Access Management Areas. *Frontiers in Marine Science* 8.

Romero, P., and O. Melo. 2021. Can a Territorial Use Right for Fisheries management make a difference for fishing communities? *Marine Policy* 124:104359.

Rustagi, D., S. Engel, and M. Kosfeld. 2010. Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science* 330(6006):961–965.

Santis, O., and C. Chávez. 2015. Quota compliance in TURFs: An experimental analysis on complementarities of formal and informal enforcement with changes in abundance. *Ecological Economics* 120:440–450.

Sen, S., and J. Raakjaer Nielsen. 1996. Fisheries co-management: a comparative analysis. *Marine Policy* 20(5):405–418.

Thöni, C., and S. Volk. 2018. Conditional cooperation: Review and refinement. *Economics Letters* 171:37–40.

Tole, L. 2010. Reforms from the ground up: a review of community-based forest management in tropical developing countries. *Environmental Management* 45(6):1312–1331.

Travers, H., T. Clements, A. Keane, and E. J. Milner-Gulland. 2011. Incentives for cooperation: The effects of institutional controls on common pool resource extraction in Cambodia. *Ecological Economics* 71:151–161.

Varughese, G., and E. Ostrom. 2001. The Contested Role of Heterogeneity in Collective Action: Some Evidence from Community Forestry in Nepal. *World Development* 29(5):747–765.

Vollan, B. 2008. Socio-ecological explanations for crowding-out effects from economic field experiments in southern Africa. *Ecological Economics* 67(4):560–573.

Wamukota, A. W., J. E. Cinner, and T. R. McClanahan. 2012. Co-management of coral reef fisheries: A critical evaluation of the literature. *Marine Policy* 36(2):481–488.

Wilkinson, C., and B. Salvat. 2012. Coastal resource degradation in the tropics: Does the tragedy of the commons apply for coral reefs, mangrove forests and seagrass beds. *Marine Pollution Bulletin* 64(6):1096–1105.

Zhu, T., M. Krott, and H. Chen. 2014. Co-management implementation in forested national reserves: Contradicting cases from China. *Forest Policy and Economics* 38:72–80.

4. SECOND PAPER. STRATEGIC HETEROGENEITY AND COOPERATION DYNAMICS IN COMMON-POOL RESOURCE DILEMMAS: EXPERIMENTAL AND SIMULATION INSIGHTS

4.1. INTRODUCTION

Social dilemmas are scenarios of interdependent decision-making characterized by the intrinsic tension between individual and collective interests, where selfish actions can lead to suboptimal outcomes from a group perspective. Theoretical models based only on the assumption of self-interested preferences predict minimal levels of cooperation, leading to the non-provision of public goods or the depletion of common-pool resources (CPRs). However, contrary to these predictions, case studies and experimental evidence have revealed conditions and scenarios that allow groups to avoid such critical outcomes and exhibit sharp cooperation despite the incentives to defect from collective action (Ostrom, 1990). Public goods are frequently created, and the tragedy of the commons can effectively be averted, even in large-scale societies where models based on reciprocal altruism or indirect reciprocity do not seem to provide a sufficient explanation. But cooperation in social dilemmas is also complex and characterized by its fragility. Repeated interactions without incentives such as rewards or punishment showcase a consistent downward spiral in cooperation levels and, while cooperation exceeds theoretical predictions initially, there is an observable decline over time (Balliet et al., 2011; Ledyard, 1995).

To explain these observed patterns of cooperation in social dilemmas, experimental economists have proposed the existence of, and tested the effects of, preference and strategic heterogeneity—specifically, the existence of behaviors that deviate from the classical economic assumption of self-interest (Charness & Rabin, 2002; Fehr & Fischbacher, 2002; Fehr & Schmidt, 1999a; Sobel, 2005). In the domain of social dilemmas and game theory, strategies represent plans of action that subjects adopt in response to the behaviors of others and to an incentive structure. These actions may align with either self-interested (e.g., selfishness) or social preferences (such as altruism, reciprocity, or inequity aversion), thereby embodying varying levels of cooperation. The existence of strategies deviating from self-interested predictions is responsible for both the higher-than-anticipated levels of cooperation and its decline in repeated interactions (Ambrus & Pathak, 2011; Cheung, 2014; Fischbacher & Gächter, 2010). This strategic heterogeneity has been consistently detected through experimental game theory, leading to the characterization of prototypical strategies such as conditional cooperation, unconditional cooperation, and free-riding, among others (Fallucchi et al., 2019; Fischbacher et al., 2001; Kocher et al., 2008; Kurzban & Houser, 2005; Rustagi et al., 2010).

Experimental scholars have employed methodologies such as the Strategy Method (Fischbacher et al., 2001) and Linear Conditional Contribution Profiles (Kurzban & Houser, 2005; van Klingeren, 2022) to assess the impact of group composition, including strategic heterogeneity, on group-level performance in social dilemmas. (see Guido et al., 2019 for a meta-analysis). Gächter & Thöni (2005) found that like-minded groups generally exhibited higher contributions than randomly assembled groups, especially in the presence of costly peer punishment. Similarly, Burlando & Guala (2005) showed that homogeneous groups of cooperators and reciprocators maintained high contributions, while groups of free-riders experienced a decline. Ones & Putterman (2007) demonstrated that strategies remained stable across rounds, and group composition strongly predicted performance. de Oliveira et al. (2015) further highlighted that the presence of free-riders not only reduces their own contributions but also lowers those of conditional cooperators. This body of literature underscores the importance of understanding how heterogeneous strategies interact to shape cooperative dynamics. However, due to the inherent limitations of experimental procedures, significant gaps remain in our understanding of the potential effects of strategic heterogeneity, particularly because it is not feasible to experimentally analyze a wide range of combinations of strategies.

Simulation techniques, such as agent-based models (ABMs), have helped bridge this gap, though models specifically tailored to this field remain relatively scarce. In a study similar to our proposal, Lucas et al. (2014) designed an ABM based on the experimental results of Fischbacher & Gächter (2010). In this model, simulated agents contributed in a 10-round public goods dilemma according to a strategy, a belief in the contributions of other agents, or a combination of both. The model encompassed a wide range of strategies, including unconditional cooperation and free-riding as unconditional strategies, and conditional cooperation and negative cooperation as conditional strategies, among others. The researchers simulated 495 possible group compositions and used a random model regression to estimate the marginal effect of increasing each type of strategy on the average group contribution. Notably, their results show that, compared to free-riders, unconditional cooperators have a greater impact on average group contributions than conditional cooperators; that contributions consistently decline over time; and that, under certain model specifications, heterogeneity in social strategies significantly reduces average group contributions. Another ABM designed to explore strategic and preference heterogeneity on cooperative dynamics was developed by Amin et al. (2018). They used a two-stage experimental procedure to categorize subjects' preferences and then used these preferences to determine the agents' contributions in a repeated public goods model, varying the prevalence of each strategic type in a heterogeneous population from 0% to 100%. Their results suggest that Nash equilibrium is only achieved with a high prevalence of free-riders (at least 40% within the population), challenging previous literature that attributes greater influence to a small number of free-riders: Fehr & Schmidt, 1999).

Although these gaps have started to close, researchers still have a limited understanding of how strategic heterogeneity specifically affects CPR management, as most studies rely on public-good games, neutrally framed instructions, and student samples. The contribution of this study is to explore the impact of different strategies and strategic profiles on group performance, focusing on how the prevalence of these strategies shapes both group cooperation and the overall dynamics within groups. We model strategies that have been robustly identified in the literature, such as free-riding, conditional cooperation, and unconditional cooperation, while also incorporating a frequently overlooked strategy type called negative cooperation, which has already been identified in experimental settings (Croson, 2007; Lucas et al., 2014). To classify and characterize these strategies, we initially conducted a CPR game experiment with small-scale fishers. Afterward, we integrate these experimental results into the design of a parsimonious agent-based model, combining the advantage of highly controlled conditions in experimental research with the flexibility and depth of simulation models for understanding social behavior. This contribution to the existing literature is grounded in findings from a lab-in-the-field experiment that involved real CPR users rather than undergraduate students. To enhance ecological validity, we implemented several measures, including a treatment involving external, non-deterrent norm enforcement and the use of contextually framed instructions. Additionally, our sample was pre-classified based on participants' real-life experience in a specific institutional arrangement for CPR governance: the co-management of benthic fishery resources. Our sample is composed of three types of user groups: fishers with high real-life cooperation in co-management, fishers with low real-life cooperation in co-management, and fishers not involved in co-management. This real-life behavior is reflected in the experiment, as these differences are also detected in our CPR game. To date and to our knowledge, this is the first attempt of modelling the effect of strategic heterogeneity informed by a contextually framed experiment performed with real CPR users and with this pre-classified sample. Employing real-world users, rather than typical laboratory subjects such as university students, enhances the external validity of the findings, allowing for conclusions that are more applicable to actual resource management scenarios. Moreover, incorporating experimental parameters derived from real-world contexts into simulation models strengthens their predictive power and provides a valuable tool for designing effective public policies for CPR management.

4.2. SUBJECT POOL AND EXPERIMENTAL PROCEDURE

Chilean benthic fisheries were unregulated and subject to an open-access regime until 1997. Due to increasing economic demands in previous decades, benthic resources became severely depleted, leading to the introduction of a new regulatory framework based on co-management, known as the Management and Exploitation Areas for Benthic Resources (MEABRs). Under this regulatory mechanism, small-scale fishers are required to form unions to apply for exclusive rights to exploit resources within these designated areas. Fisher unions, with the assistance of specialized consultants, develop management plans and oversee the monitoring of exploitation agreements within the union, while state agencies provide external enforcement against norm violators. In contrast, fishers who choose not to participate in MEABRs exploit benthic resources individually and outside these designated areas. Some may attempt illicit access to these areas, undermining the ecological and economic outcomes for the fishing unions.

We conducted a lab-in-the-field repeated CPR experiment with 85 small-scale benthic fishers from Chile. Of this sample, 55 participants were affiliated with six MEABRs and belonged to two types of user groups: high-performance unions (HP), characterized by high levels of compliance in benthic fishery co-management, and low-performance unions (LP), marked by low levels of cooperation in co-management. This classification (HP and LP) was constructed on the basis of data collected by Marín et al. (2012) and Gelcich et al. (2013), and accounted for variables such as compliance with union norms and rules, co-management performance measures (assessed by the National Fisheries and Aquaculture Service), and ecological outcomes (measured through the evolution of the Total Allowable Catch and biodiversity levels). The remaining 30 participants were not affiliated with any union and thus extracted resources outside of MEABRs. To address the substantial biological and sociocultural diversity along the Chilean coastline, we intentionally restricted our subject pool to a specific area on the central coast, a culturally homogeneous population inhabiting a shared geographical and ecological context. All experimental subjects lived in coastal settlements of two Chilean administrative areas (Valparaíso and O'Higgins regions) and shared a small-scale fisher's culture which combines traditional and modern methods, and the same target economic species. Therefore, only organizational features distinguish HP and LP user groups, as all other geographical, ecological and cultural characteristics are shared between subjects. In summary, the study's subject pool can be categorized into three distinct user groups: fishers affiliated with high-performance unions (HP, n = 30), fishers affiliated with low-performance unions (LP, n = 25), and non-unionized fishers (NU, n = 30). All experimental sessions were conducted in 2012 in the coastal settlements where participants lived.

The experiment consisted of a repeated common-pool resource game with a within-subjects design, featuring two treatments: an unenforced fishing quota treatment in rounds 1 to 10 (hereafter referred to as the unenforced treatment) and an enforced fishing quota treatment in rounds 11 to 20 (the enforced treatment). Participants were assembled into groups of five individuals belonging to the same fishing union, while non-unionized fisher were just assembled between them. These groups were randomly formed and remained fixed throughout the experimental session. The experiment included anonymization protocols to minimize the impact of third-party observation and reputational concerns. Participants were not allowed to communicate with each other, and payments were kept confidential and disbursed at the end of each session. Furthermore, the instructions were carefully framed to reflect real-life situations commonly encountered by participants in their daily economic activities. In every round of the game, subjects were told that they possessed an initial endowment of 100 units of a local benthic fishery resource known as 'Loco' (*Concholepas concholepas*), which was referred to in the experiment instructions as the individual allowed quota. Additionally, in each round, subjects simultaneously decide the extent of overharvest, ranging from 0 to 50 units of Loco. This overharvest, denoted as $y_{it} \in \{0,...,50\}$, represents the potential extra units of Loco extracted by subject *i* in round *t*. Each overharvested unit contributed to the final individual payoff of *i*, but also caused a half-unit loss of Loco for every other member of the group, mimicking the real-life collective negative externality associated with resource overexploitation.

In the unenforced treatment (rounds 1 to 10), the basic experimental setup allowed players to independently decide the extent of over-extraction in each round, with no enforcement mechanism to ensure compliance with the individual quota. At the end of each round, participants received information on i) the average extraction of their group, ii) their individual payoff for that round, and iii) their individual losses due to the group's over-extracted Locos. No data were provided regarding the individual over-extraction levels of teammates, preserving the privacy of each player's behavior throughout the experiment. In the enforced treatment (rounds 11 to 20), we introduced a punitive but non-deterrent external enforcement mechanism designed to resemble the institutional monitoring and sanctioning practiced by government agencies. Under this treatment, subjects still possessed 100 units of Loco and they had to decide how many extra units overharvest per round. However, at the end of each round, two subjects per group were randomly selected for inspection. If the inspected subject overharvested (exceeding their individual quota, i.e., if $y_{it} > 0$), they lost the entire payoff for that specific round. In both treatments, each acquired 'Loco' unit was assigned a value of $10 CLP (10 Chilean pesos) and, thus, individual payoffs per round can be expressed as:

Equation 1

$$\pi_{it} = \begin{cases} \$10(100 + y_{it} - \frac{1}{2}\sum_{j-i} y_t), & if\ t \leq 10 \\ \$10(100 + y_{it} - \frac{1}{2}\sum_{j-i} y_t), & if\ y_i = 0\ and\ t > 10 \\ \frac{3}{5}\$10(100 + y_{it} - \frac{1}{2}\sum_{j-i} y_t), & if\ y_i > 0\ and\ t > 10 \end{cases}$$

The social optimum in this context is achieved when all participants choose not to overharvest any additional units beyond their individual allowed quota (i.e., when $y_{it} = 0$ for all *i* and *t*). This outcome maximizes the collective welfare of the group by preserving the resource and avoiding the negative externalities associated with overharvesting. If all participants comply fully, the group's total extraction remains within sustainable limits, preventing resource depletion and ensuring that everyone receives the maximum possible payoff per round. The Nash equilibrium, however, differs from the social optimum. In the absence of enforcement (unenforced treatment, rounds 1 to 10), each participant has an incentive to overharvest since the individual benefit from overharvesting (gaining an additional $10 CLP per unit overharvested) outweighs the individual cost ($5 CLP loss per unit for each other group member). This incentive leads to a situation in which, in equilibrium, each participant overharvests as much as possible (i.e., $y_{it} = 50$). In the enforced treatment (rounds 11 to 20), the introduction of an external non-deterrent enforcement mechanism changes the incentive structure. If a participant is caught overharvesting, they lose their entire payoff for that round. However, the enforcement is non-deterrent and, consequently, some individuals may still choose to overharvest, anticipating that the potential reward outweighs the risk of punishment. The Nash equilibrium in this scenario is more complex, as it depends on each participant's risk tolerance and belief about others' behavior. More specifically, a subject will refrain from over-extracting only if she expects her teammates to over-extract 50 or fewer units of loco in round t; if she anticipates that this condition will not be met, she may find no incentive to limit her own over-extraction.

To analyze decisions within the context of cooperation, we transformed each subject's overharvest level into an individual normalized compliance value per round ($[50 - y_{it}] / 50$). Therefore, y_{it} is hereafter the normalized compliance value of subject *i* in round *t*, ranging continuously from 0 (total overharvest) to 1 (total compliance, i.e., no overharvest at all). We made this transformation considering that much of the literature on social dilemmas is framed in terms of cooperation. Additionally, the instructions and parameters of the experiment allow overharvesting to be interpreted as a non-cooperative behavior, thereby justifying the validity of this transformation.

4.3. AGENT-BASED MODEL OF REPEATED CPR DILEMMA

The agent-based model was developed to parsimoniously simulate this repeated social-dilemma and to explore the effects of varying proportions of strategic types on group cooperative dynamics. This section outlines the strategy classification process and subsequently offers a detailed description of the ABM, including its features, parameters, and the metrics used to address our research questions. In appendix 1 we provide a more detailed description of the model following the ODD+D protocol proposed by Grimm et al. (2006, 2010).

4.3.1. Strategy classification method

We categorize the strategies of all subjects based on their experimental responses in each treatment, assuming that these strategies are stable over time (Kurzban & Houser, 2005; Ones & Putterman, 2007), while also recognizing that they may change in response to shifts in the incentive structure, such as those introduced in the enforcement treatment (Rodriguez-Sickert et al., 2008). As a result, subjects have two classifications, and participants could either maintain or change their strategy between treatments. We identified four possible strategies that subjects could adopt in either treatment: conditional cooperation (where compliance increases monotonically with the group's mean compliance), unconditional cooperation (consistently cooperating at high levels), negative cooperation (where compliance decreases monotonically with the group's mean compliance), and free-riding (characterized by consistently low or no cooperation, regardless of others' behaviors). To classify strategies, we estimated each subject's Linear Conditional Contribution Profile (LCCP), based on:

Equation 2

$$y_{i,t} = \alpha_{i1} + \alpha_{i2}DR + \beta_{i1}\bar{x}_{-i,t-1} + \beta_{i2}\bar{x}_{-i,t-1}DR + \varepsilon_{i,t}$$

Where $y_{i,t}$ is the compliance of subject *i* in round *t*; DR is a dummy variable that takes the unenforced treatment as the reference category; $\bar{x}_{-i,t-1}$ is the mean compliance of *i*'s partners in *t-1*; and $\varepsilon_{i,t}$ is an error term. As $\bar{x}_{-i,t-1}$ was needed, rounds 1 and 11 were dropped from regressions since there is no *t-1* in round 1, and *t-1* in round 11 has a trend of the unenforced treatment and could have distorted our results. We obtained these parameters using Constrained Least Squares, a method that imposes restrictions on the minimization of the sum of squared residuals in a linear model. We constrained the parameters to lie within the set [0, 1] to guarantee that LCCPs consistently stayed within the feasible region of the model. Utilizing these LCCPs parameters, we classified subjects' strategies in each treatment based on the following criteria: a free-rider is a subject whose LCCP is consistently lower than half of the total possible compliance (i.e., their expected compliance is consistently < 0.5 at all levels of others' contributions). This means that a subject is classified as free-rider (FR) if:

- In the unenforced treatment: $\hat{\alpha}_{i1} < 0.5$ and $\hat{\alpha}_{i1} + \hat{\beta}_{i1} < 0.5$
- In the enforced treatment: $\hat{\alpha}_{i1} + \hat{\alpha}_{i2} < 0.5$ and $\hat{\alpha}_{i1} + \hat{\alpha}_{i2} + \hat{\beta}_{i1} + \hat{\beta}_{i2} < 0.5$

A conditional cooperator (CC) complies below 0.5 when their teammates' compliance is set to zero (i.e., the intercept of their LCCP is < 0.5), but respond to their group behavior with an expected compliance above 0.5. This means that a subject is classified as conditional cooperator if:

- In the unenforced treatment: $\hat{\alpha}_{i1} < 0.5$ and $\hat{\alpha}_{i1} + \hat{\beta}_{i1} > 0.5$
- In the enforced treatment: $\hat{\alpha}_{i1} + \hat{\alpha}_{i2} < 0.5$ and $\hat{\alpha}_{i1} + \hat{\alpha}_{i2} + \hat{\beta}_{i1} + \hat{\beta}_{i2} > 0.5$

A negative cooperator (NC) complies above 0.5 when their teammates' compliance is set to zero (i.e., their LCCP intercept is > 0.5), but respond to their group behavior with an expected compliance below 0.5. This means that a subject is classified as negative cooperator if:

- In the unenforced treatment: $\widehat{\alpha}_{i1} > 0.5$ and $\widehat{\alpha}_{i1} + \widehat{\beta}_{i1} < 0.5$
- In the enforced treatment: $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2} > 0.5$ and $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2} + \widehat{\beta}_{i1} + \widehat{\beta}_{i2} < 0.5$

And finally, an unconditional cooperator (UC) is a subject who always display an LCCP above half of the total possible compliance (i.e., their expected compliance is always > 0.5 at all levels of others' contributions). This means that a subject is classified as unconditional cooperator if:

- In the unenforced treatment: $\widehat{\alpha}_{i1} > 0.5$ and $\widehat{\alpha}_{i1} + \widehat{\beta}_{i1} > 0.5$
- In the enforced treatment: $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2} > 0.5$ and $\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2} + \widehat{\beta}_{i1} + \widehat{\beta}_{i2} > 0.5$

Due to issues with variance in the independent variables, we had to exclude seven participants (five from HP and two from LP unions). As a result, the final sample sizes were set at HP, n = 23; LP, n = 24; and NU, n = 30. Using this classification method, we estimated the proportion of strategies for each user group in each treatment (hereafter referred to as the strategic profile), as shown in Tables 1 to 3. Additionally, Table 4 provides the strategic profile for the entire sample.

Table 1. Percentage of strategies in both treatments in high-performance user groups.

	Unenforced treatment	Enforced treatment	Percentage difference
Unconditional cooperators	43.48%	69.57%	26.09%
Conditional cooperators	39.13%	17.39%	-21.74%
Negative cooperators	8.70%	4.35%	-4.35%
Free riders	8.70%	8.70%	0%
Last round mean cooperation	0.612	0.863	
Mean cooperation across rounds	0.715	0.838	

Table 2. Percentage of strategies in both treatments in low-performance user groups.

	Unenforced treatment	Enforced treatment	Percentage difference
Unconditional cooperators	4.17%	25%	20.83%
Conditional cooperators	41.67%	45.83%	4.16%
Negative cooperators	4.17%	4.17%	0%
Free riders	50%	25%	-25%
Last round mean cooperation	0.04	0.446	
Mean cooperation across rounds	0.24	0.578	

Table 3. Percentage of strategies in both treatments in non-unionized user groups.

	Unenforced treatment	Enforced treatment	Percentage difference
Unconditional cooperators	0%	3.33%	3.33%
Conditional cooperators	16.67%	23.33%	6.66%
Negative cooperators	6.67%	20%	13.33%
Free riders	76.67%	53.33%	-23.34%
Last round mean cooperation	0.103	0.141	
Mean cooperation across rounds	0.198	0.288	

Table 4. Percentage of strategies in both treatments in the whole sample.

	Unenforced treatment	Enforced treatment	Percentage difference
Unconditional cooperators	14.29%	29.87%	15.58%
Conditional cooperators	31.17%	28.57%	-2.6%
Negative cooperators	6.49%	10.39%	3.9%
Free riders	48.05%	31.17%	-16.88%
Last round mean cooperation	0.25	0.48	
Mean cooperation across rounds	0.38	0.56	

We evaluated the consistency of this classification method across both treatments by applying a new set of Constrained Least Squares to test whether each classification follows a consistent functional form. In this process, we pooled the data of subjects of the same strategy and introduced dummy variables for the enforced treatment. If the estimated parameters of these dummy variables are not statistically significant, we can infer that the classification maintains a consistent functional form across both treatments, allowing us to use the pooled data in the ABM. In each constrained model, we used the observations of agents classified under each strategy to estimate:

Equation 3

$$y_{i,t} = \beta_1 + \gamma_1 DR + \beta_2 \overline{\mathrm{x}}_{-i,t-1} + \gamma_2 \bar{x}_{-i,t-1} DR + \varepsilon_{i,t}$$

Where: y_{it} is the normalized compliance of individual *i* in time *t*; β_1 is the level of unconditional cooperation in the unenforced treatment; γ_1 is the change of unconditional cooperation in the enforced treatment; β_2 is the level of conditional cooperation in the unenforced treatment; and γ_2 is the change of the conditional cooperation in the enforced treatment relative to the unenforced treatment. If the classification is consistent, parameters $\hat{\gamma}_1$ and $\hat{\gamma}_2$ should not be statistically significant. The results shown in table 5 indicate that as expected, both parameters are not statistically significant at the 95% level in any of the four models:

Table 5. P-values of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ in each model

	P-value of $\hat{\gamma}_1$	P-value of $\hat{\gamma}_2$
Unconditional cooperators model	0.0756	0.1650
Conditional cooperators model	0.8558	0.7399
Negative cooperators model	0.9110	0.2053
Free riders' model	0.6262	0.7788

As the classification of strategies is consistent and robust across treatments, we used the pooled data of subjects displaying the same strategy to estimate what we will call prototypical behavioral models, i.e., models that generically represent each of the four strategies. This time, we ran Tobit regressions for each of the strategies with left and right censoring (at 0 and 1, respectively), as follows:

Equation 4

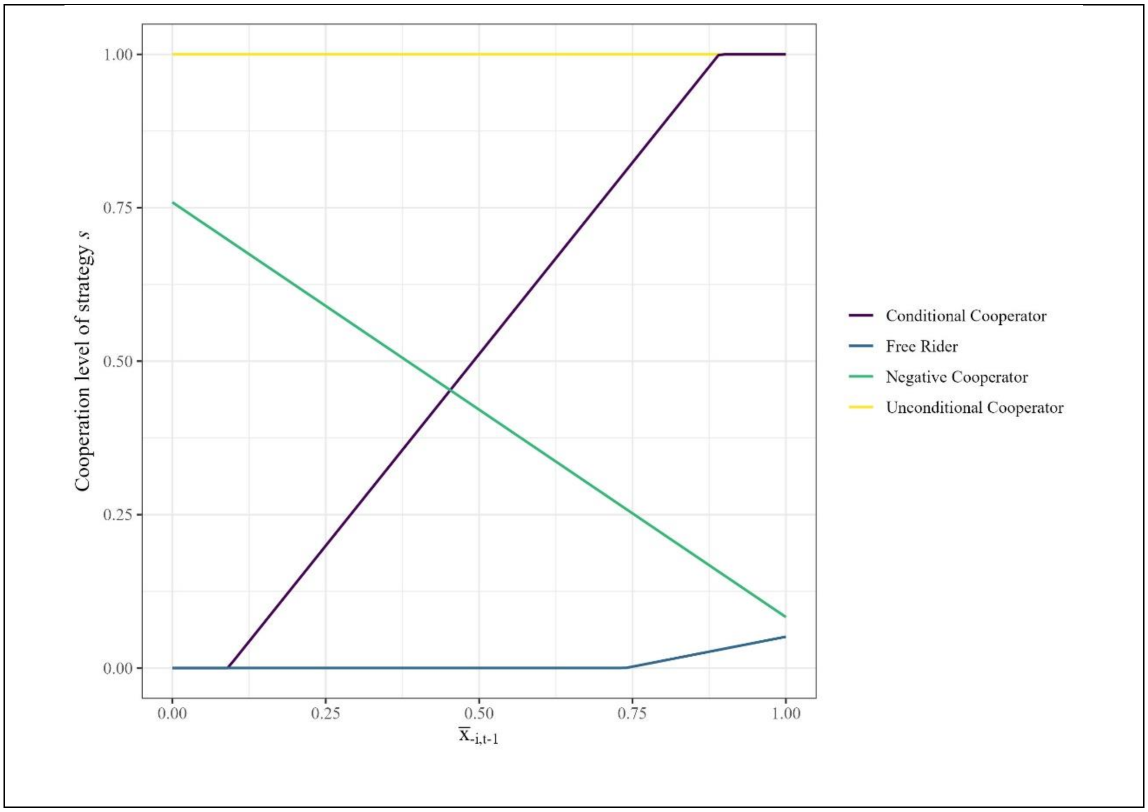
$$y_{i,t} = \beta_1 + \beta_2 \bar{x}_{-i,t-1} + \varepsilon_{i,t}$$

Where: y_{it} is the compliance of individual *i* in time *t*; β_1, the intercept, represents the level of unconditional cooperation; β_2, the slope, represents the level of conditional cooperation; and ε_{it} is an error term. Since we were able to use all observations of agents classified under each strategy, in this new set of models there are no parameters associated with dummy variables for the enforced treatment. The absence of such parameters indicates that these models do not require separate treatment-specific coefficients. Table 6 presents the results of the Tobit estimation for all prototypical models, including all three relevant parameters to simulate agents' behavior:

Table 6. Prototypical models (TOBIT)

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\sigma}_u$ (residual standard deviation)
Unconditional cooperators model	1.011	0.344	0.595
Conditional cooperators model	-0.113	1.249	0.644
Negative cooperators model	0.759	-0.676	0.748
Free riders' model	-0.144	0.195	0.470

Figure 1. Prototypical models of strategies *s*, without residual standard deviations. Cooperation levels are truncated within the set [0, 1].



4.3.2. ABM description

The model simulates the behavior of N agents (n=1, 2, 3...N) interacting over T rounds (t=1, 2, 3...T) in a social dilemma that resembles our experimental setting. Agents belong to a strategy $s \in$ {FR, CC, NC, UC} and are assigned to its prototypical model to simultaneously contribute a compliance level of y_{it}. At the beginning of each round, we update the information of the average contributions of the rest of the population in the previous round ($\bar{x}_{-i,t-1}$) and use it to estimate agent *i*'s current contribution. For the case of round 1, it was not possible to use the prototypical behavior models as round 1 was excluded from those regressions. Thus, for $y_{i,s,t=1}$ we sample a value within the empirical distribution estimated by non-parametric bootstrapping of the experimental contributions in round 1 of strategy s. For subsequent rounds (t>1), agents contribute according to their prototypical model, as follows:

Equation 5

$$y_{i,s,t>1} = max(min(\hat{\beta}_{1s} + \hat{\beta}_{2s}\bar{x}_{-i,t-1} + \hat{u}_{s,t}, 1), 0)$$

Where $y_{ist>1}$ is the contribution of agent *i* in round t>1, while $\hat{\beta}_{1s}$ and $\hat{\beta}_{2s}$ are the estimated parameters of that prototypical behavior model. To introduce variability, we include an error term $\hat{u}_{st},$ which is a normally distributed random variable with mean 0 and standard deviation equal to $\hat{\sigma}_s$ (the residual standard deviation of prototypical model of s). This random error accounts for possible variability between subjects following the same strategy, as well as the influence of uncertainty or trembling hands in individual decision-making during each period. To prevent non-feasible contributions, we truncate agents' responses to the range [0, 1], as it was not possible to cooperate beyond or below these bounds in the experimental procedure.

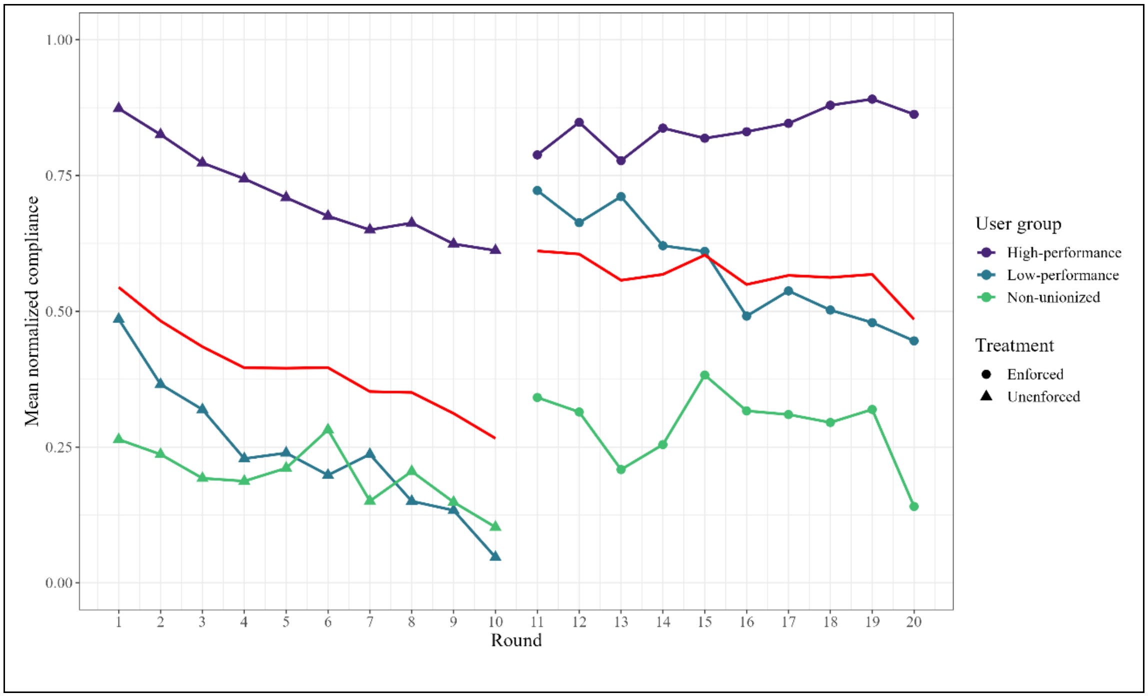
To investigate how varying the percentage of each strategy affects group performance and the stability of cooperation, we ran simulations with different strategic compositions. We studied one strategy at a time, increasing its prevalence from 0% to 100% in 1% increments. This method allows us to estimate how the percentage increase in the prevalence of a strategy s (the focal strategy) affects group behavior. Agents not assigned to s were assigned to non-focal strategies using the empirical distribution inferred from experimental data. Let x% be the prevalence of the studied strategy s in a given simulation, which we set exogenously. Let b%, c%, and d% be the proportions of non-focal strategies in the subject pool. We allocated the remaining (100-x)% of simulated agents to strategies B, C, and D in the proportions [b/(b+c+d)](100-x), [c/(b+c+d)](100-x), and [d/(b+c+d)](100-x), respectively. For example, if we aim to study the effects of 20% of FRs in the HP user group under the unenforced treatment, we take the empirical distribution of the non-focal strategies in that strategic profile (43.48% of UCs, 39.13% of CCs, and 8.7% of NCs). Since these proportions sum to 91.31%, we allocate the remaining 80% of the simulated agents according to this distribution among them: 38.09% of UCs ([43.48/91.31]*80), 34.28% of CCs ([39.13/91.31]*80), and 7.62% of NCs ([8.7/91.31]*80). We ran 500 simulations for each 1% increase in the representation of each strategy and evaluated the model's outcomes using two metrics: group outcomes, measured as the mean cooperation level of the N agents in the final round ($\bar{y}_{t=T}$) and the cooperative stability of groups, defined as the percentage of simulations where the mean cooperation in the final round was equal to or higher than the initial mean.

4.4. RESULTS

4.4.1. Experimental results

In Figure 2 we show the experimental findings from the CPR lab-in-the-field, particularly the mean normalized compliance across rounds of each user group. This result suggests that only the high-performance unions in the enforced treatment are capable of sustaining cooperation across rounds; that non-unionized fishers do not seem to be affected by the external punishment treatment; and that low-performance unions are capable of rising cooperation in the first round of the enforced compared to the first round of unenforced treatment, but that are also incapable of preventing its erosion across rounds.

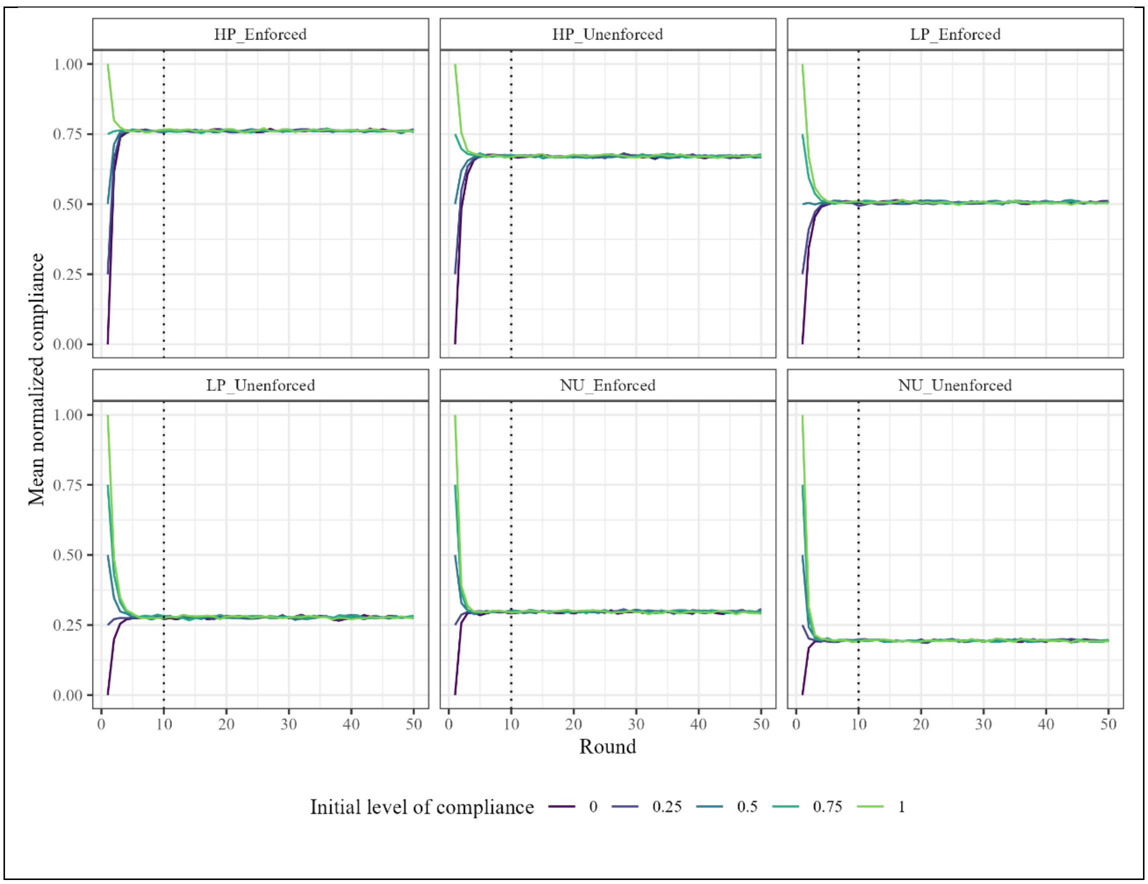
Figure 2. Mean normalized compliance across round by user group. In red: overall mean normalized compliance.



4.4.2. ABM Results

Before presenting our main results, we first test whether different initial contributions in round 1 and fixed proportions led to single or multiple equilibria. This step is crucial, as all subsequent analyses rely on the assumption that strategic heterogeneity determines a group's cooperative path independently of their initial cooperation levels. To test this, we simulated each of the six proportions from the experimental results with varying initial levels of compliance in round 1 (running 1,000 simulations per initial level). The proportions were held constant over 50 rounds to examine whether fixed proportions with different initial cooperation levels would lead to unique or multiple equilibria. Figure 3 shows the mean level of cooperation per round across different initial contribution levels, providing evidence in favor of our assumption by demonstrating that unique equilibria were reached in all experimental conditions, often before round 10.

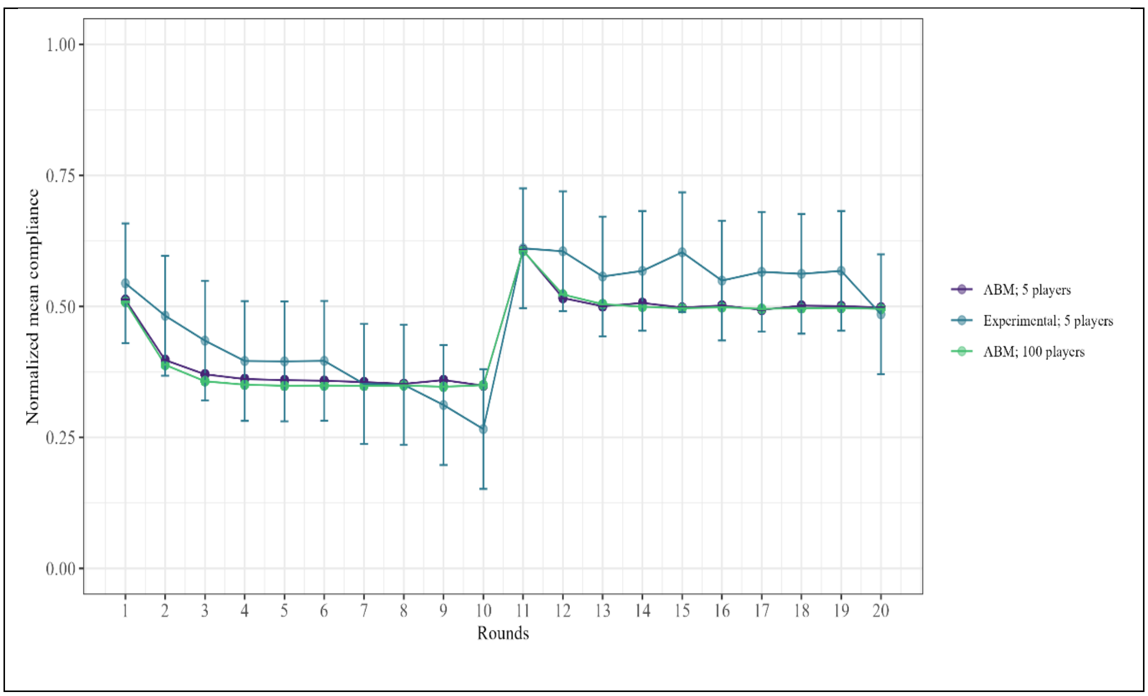
Figure 3. Different initial levels of compliance and single equilibriums.



HP: High-performance strategic profile; **LP**: Low-performance strategic profile; **NU**: Non-unionized strategic profile.

To evaluate the predictive power of our simulation model, we compared the mean levels of normalized compliance between the experimental results and two settings of the ABM: one with groups of five agents (resembling the experimental setting) and one with 100 agents. We used the whole sample strategic profile (presented in table 4) and simulated each round 1000 times for each of the ABM group settings. In Figure 4, we show the results of both ABM settings alongside the experimental mean levels of compliance for each round, demonstrating that both settings closely predict the experimental data.

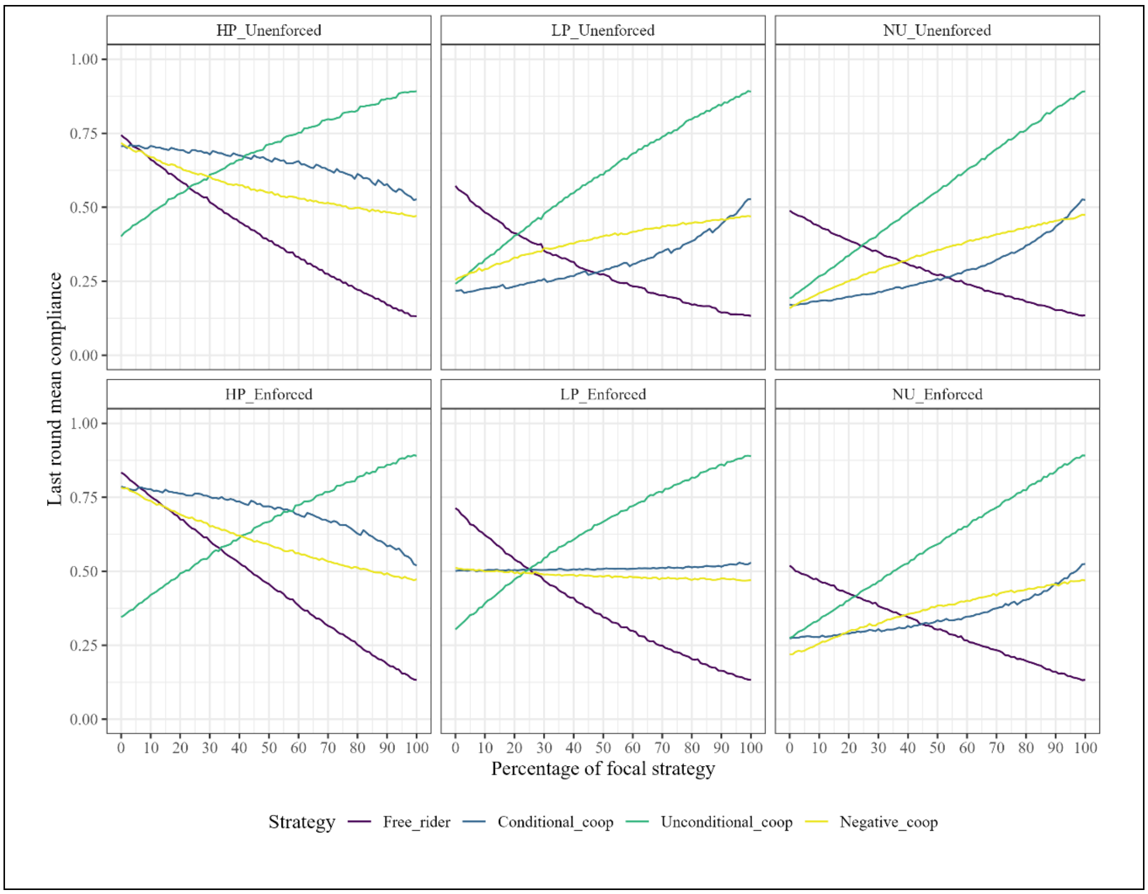
Figure 4. Experimental and ABMs mean compliance across rounds.



Experimental results with standard deviation results after non-parametric block bootstrapping.

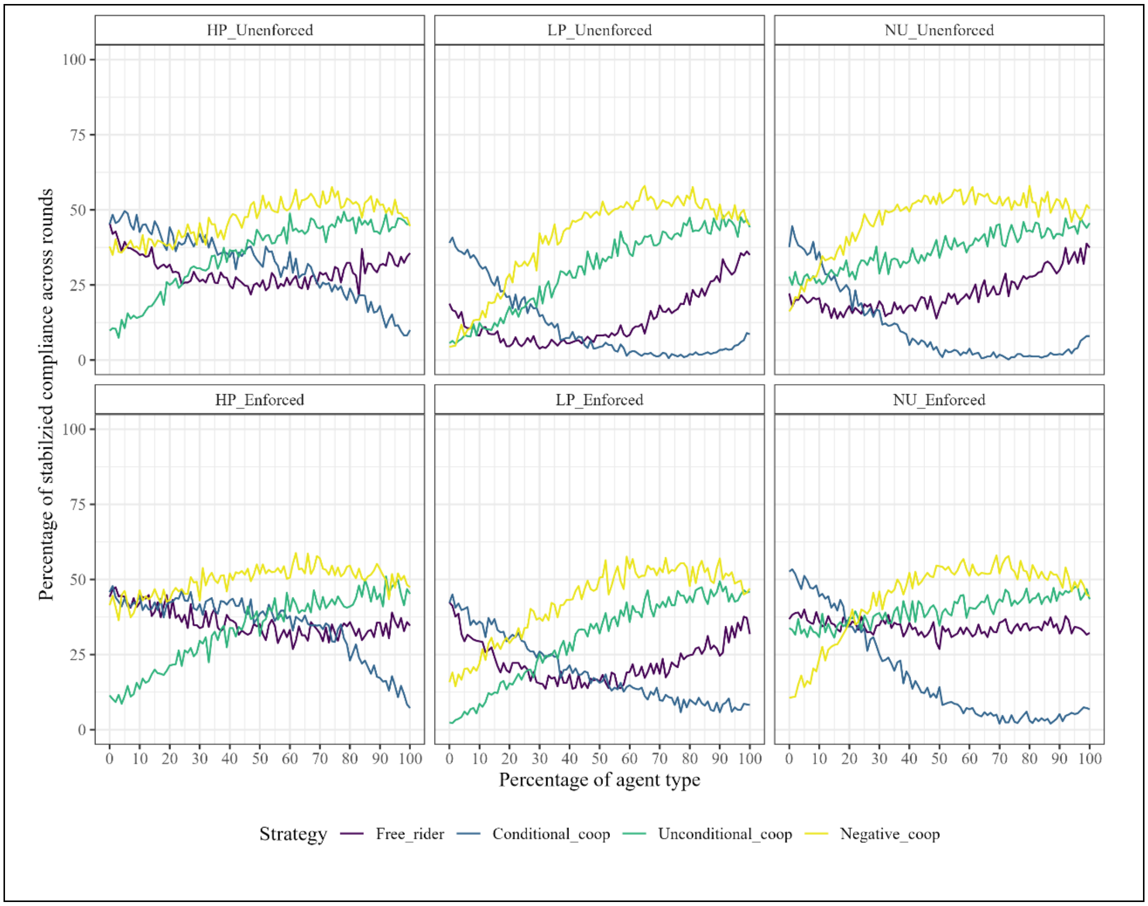
One of the main results of this study is illustrated in Figure 5, where the mean compliance in the last round is shown for each percentage of prevalence across all experimental strategic profiles. This result shows that in the best strategic profile scenario (HP in the enforced treatment), at least 20% of unconditional cooperators (UCs) are required to achieve more than 50% mean compliance in the last round. However, even in this optimal strategic profile scenario, the presence of around 45% of free-riders (FRs) is enough to bring mean compliance below 50% by the end of the game. Although this may seem like a large proportion of FRs, other strategic profiles require a much smaller share of free-riders to reach this same outcome. In both non-unionized strategic profiles, for instance, no FRs are needed, whereas for the LP user groups between 10% and 25% of FRs are sufficient for the same result, respectively. These findings indicate that, depending on the profile of other strategies, even a minor share of FRs can lead to a significant reduction in overall compliance. The effects of increasing the prevalence of both conditional and negative cooperators show similarities. For the case of both treatments in the high-performing user groups, a larger share of these strategies is associated with a decrease in mean compliance in the last round. For both treatments in the non-unionized strategic profile and for the low-performing user groups under the unenforced treatment, a larger share of both strategies is associated with an increase in mean compliance in the last round. These results suggest that increasing both strategies shows opposite effects depending on the strategic profile of the user group. For the low-performing strategic profile in the enforced treatment, an increase in both strategies does not seem to affect mean compliance in the last round.

Figure 5. Mean compliance in last round.



In Figure 6, we plot the effects of increasing the prevalence of each strategy (x-axis) on the percentage of stabilized groups (y-axis), revealing several key patterns. First, across all strategic profiles, an increase in the percentage of conditional cooperators (CCs) results in a decrease in the proportion of groups that stabilize cooperation. In contrast, the presence of some percentage of free-riders (FRs) is more effective at stabilizing cooperation than the presence of conditional cooperators in all profiles. Notably, in the LP Enforced strategic profile, the increase in FRs follows a convex pattern, with the lowest level of stabilized groups occurring at around 50% of FRs. Beyond this point, further increases in FR prevalence led to a higher percentage of groups achieving stable cooperation. Additionally, in all strategic profiles except for the non-unionized groups, NCs produce the most favorable outcomes for stabilizing cooperation, even outperforming UCs. In the non-unionized profiles, NCs also perform better, but only after reaching certain threshold levels of prevalence. Across all strategic profiles, the highest percentages of stabilized groups are observed when the prevalence of NCs is between 50% and 70%.

Figure 6. Percentage of groups that stabilized cooperation.



While the separate analysis of the mean compliance in the final round and the proportion of stabilized groups provides valuable insights into the dynamics of cooperation, these measures can be misleading when considered in isolation. For instance, as shown in the stability results, a high prevalence of FRs is associated with greater stability. However, this "stability" may represent the stabilization of cooperation at very low levels, far from desirable outcomes. If we assume that both high cooperation and stability are crucial for groups managing CPRs then it is necessary to consider these two dimensions in an integrated manner. To address this, we introduce two new indicators -**Combined Cooperation-Stability Index (CCSI)** and **Cooperation Efficiency (CE)** - which account for both the level of cooperation and its consistency across rounds. The CCSI provides a flexible, weighted combination of both dimensions, where higher values indicate favorable outcomes in both cooperation and stability:

$$CCSI_{s,p} = \alpha\left(Last\ round\ mean\ compliance_{s,p}\right) + (1-\alpha)Stability_{s,p}/100$$

Subscripts *s* and *p* account for the strategy *s* ∈ {FR, CC, NC, UC} and for the prevalence *p* ∈ {0, 1,...100}, respectively, and α is a weight that can be adjusted to place more importance on either cooperation or stability. On the other hand, the Cooperation Efficiency (CE) estimator penalizes low stability more directly by reducing the overall cooperation score when stability is poor, ensuring that high cooperation is only considered effective when it is also consistent:

$$CE_{s,p} = \frac{Last\ round\ mean\ compliance_{s,p}}{1 + (1 - Stability_{s,p}/100)}$$

In figures 7 and 8 we plot both CCSI (with α=0.5) and CE across strategic profiles:

Figure 7. CCSI.

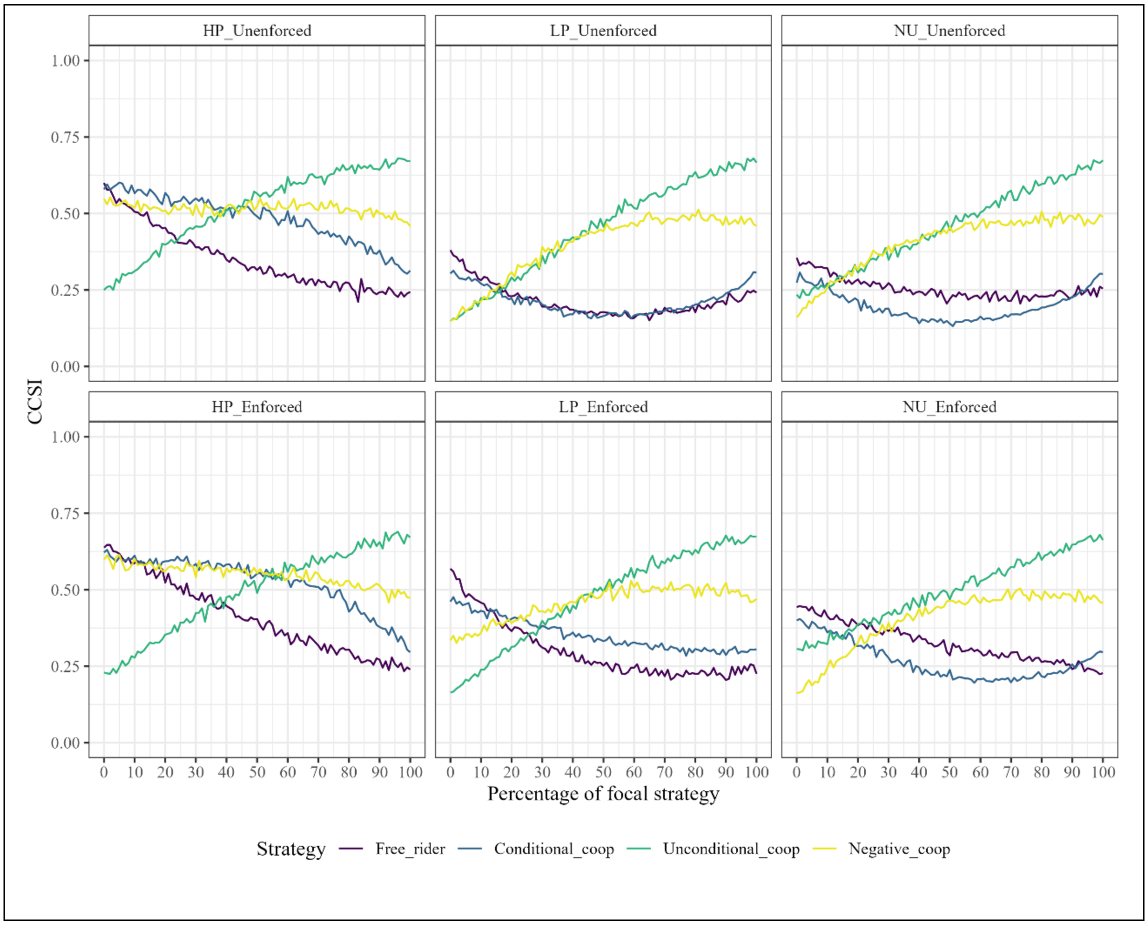
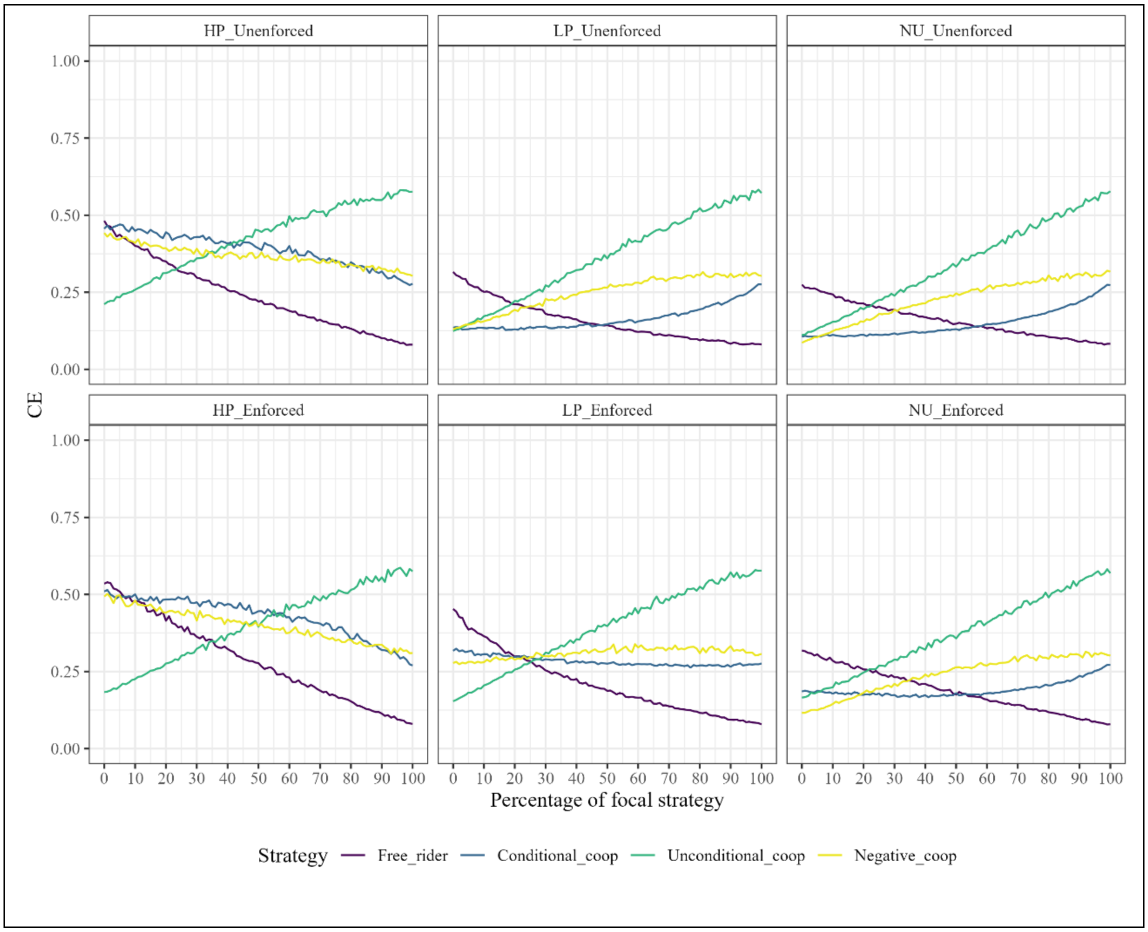


Figure 8. CE.



Both CCSI and CE show that across all strategic profiles, increasing the prevalence of FR has a negative effect on group outcomes, while increasing the prevalence of UC has a positive effect. This is reflected in the sharp decrease in both CCSI and CE values as the proportion of FR increases, indicating that while free-riders may stabilize cooperation, they do so at very low levels. In contrast, UC emerges as the most effective strategy for increasing and maintaining cooperation at high levels. On the other hand, conditional strategies (CC and NC) show more complex and moderate effects, with variations that do not significantly improve cooperation or stability compared to UC. More interestingly, in both CCSI and CE there are also opposite effects of increasing the prevalence of both strategies in different strategic profiles.

4.5. DISCUSSION

Our study explores the critical role that strategic heterogeneity plays in determining the overall outcome and the stability of cooperation in social dilemmas, specifically within the context of CPR management. The experimental results demonstrate that HP user groups are more capable of sustaining cooperation, particularly under the external enforcement treatment, while LP and NU user groups struggle to sustain compliance over time. This suggests that the organizational structure and initial strategic composition of a group are crucial determinants of sustained cooperative behavior, even in the presence of external monitoring and enforcement. Results of the ABM further reinforce the importance of strategic composition. They indicate that even in favorable conditions—such as HP groups under enforced treatments—a substantial proportion of UCs (around 20%) is required to achieve cooperation levels above 50%. However, our study also reveals the fragile nature of cooperation in the presence of free-riders, as even a small fraction of FRs can lead to a significant decline in group cooperation (especially in LP and NU user groups).

Interestingly, the role of NCs emerged as a critical factor in stabilizing cooperation. In most strategic profiles, NCs were more effective at stabilizing cooperation than both CCs and FRs, particularly when their prevalence ranged from 50% to 70%. This finding diverges from traditional views of cooperation dynamics, which often emphasize the detrimental role of non-cooperative strategies. Instead, our results suggest that negative cooperation, when combined with other strategies, can paradoxically support group stability under certain conditions. When we estimate metrics that integrate both final cooperation levels and stability, the effect of NCs is surpassed by that of UCs, but still yields better results than CCs. NCs can be understood as economic altruists, whose behavior mirrors the dynamics of pure altruism as defined in economic theory (Croson, 2007). In this framework, the well-being of others is integrated into an individual's utility function, but crucially, the contributions of others to the public good are seen as substitutes for one's own contributions (Lucas et al., 2014). As a result, when other members of the group contribute more to the collective good, NCs reduce their own contributions; on the contrary, when other members of the group contribute at low levels, NCs increase their cooperation, counterbalancing the actions of FRs and CCs (and the effect of their interaction), preventing both the erosion of cooperation and extreme fluctuations in cooperation levels. This behavior reflects their preference for ensuring the public good is maintained without feeling the need to personally contribute when others are already doing so. While NCs may not directly maximize cooperation levels, their role in moderating extreme behaviors within the group can paradoxically contribute to maintaining stability in heterogeneous strategic settings.

CCs display a reciprocating strategy, adjusting their level of cooperation based on the observed behavior of others. CCs cooperate when they perceive others are doing the same but reduce their contributions when faced with defection. Contrary to NCs and their stabilizing effect, CCs strategy introduce amplifying feedback into group behavior: when cooperation is high, CCs amplify it by increasing their own contributions, but when cooperation diminishes, CCs can accelerate the decline. As a result, CCs play a pivotal role in determining the trajectory of group cooperation. In groups with a strong cooperative foundation, CCs help reinforce collective action by reciprocating the positive behavior of others. However, in environments where free-riders are prevalent, CCs may quickly reduce their cooperation, leading to a rapid collapse in group contributions as an outcome of this interaction (as hypothesized by Fischbacher & Gächter, 2010). Figures 7 and 8 show that, in strategic profiles with high relative prevalence of FRs (in both NU user groups), only increasing the prevalence of CCs above 50% of the sample has positive effects on group outcomes. This suggests that, below that threshold, the interaction of CCs and FRs hinders high and sustained levels of cooperation, and only when the ratio of CCs relative to FRs is sufficiently large, there is a positive effect on group behavior. Therefore, the influence of CCs on group dynamics is highly context-dependent, with their potential to either enhance or erode cooperation contingent on the strategic composition of the group.

UCs are the cornerstone of sustained high-level cooperation in groups. These individuals consistently contribute to the collective good regardless of the behavior of others, providing a stable foundation upon which cooperative norms can be built. Their steady commitment to cooperation may boost optimal group-level outcomes, especially if there are sufficiently large interactions with CCs who may initially hesitate to contribute. The presence of UCs increases the likelihood that cooperative behavior will flourish within the group, even in environments with strategic heterogeneity. Furthermore, UCs play a critical role in protecting the group from the detrimental effects of FRs. By maintaining a constant level of contribution, UCs prevent the complete collapse of cooperation and help ensure that the group achieves higher levels of collective benefit over time.

FRs, in contrast, pose a significant threat to the maintenance of group cooperation. Their strategy revolves around maximizing individual gain by exploiting shared resources without contributing to the collective effort. The presence of FRs tends to erode group cooperation, particularly when they reach a critical threshold within the group. The strategic behavior of FRs discourages others from contributing, as individuals are less likely to cooperate if they perceive that their efforts are being taken advantage of by non-cooperators. Over time, this dynamic leads to groups dragging cooperation at minimum levels in a situation where mutual defection becomes the norm. However, FRs can also stabilize cooperation at very low levels, where cooperation persists, albeit in a diminished form, as we can see in the stability plot and both CCSI and CE plots.

One of the most surprising findings from the simulation is the effect of the percentage increase of CC and NC on group performance (whether in terms of final cooperation levels, CCSI, or CE). The results show that the effect of increasing these strategies diverges depending on the strategic profile. Unlike unconditional strategies, the prevalence of these strategies can either positively or negatively affect group performance, depending on the group's overall strategic composition. This allows us to extend and contextualize the previous findings of Lucas et al. (2014) and Amin et al. (2018), in the sense that it enables us to understand them as outcomes dependent on the specific strategic profiles used in their studies. Our contribution to this literature is to provide a broader and more refined framework of strategic profiles found in real CPR user groups, within which the NC and CC context-dependent effect emerges —effects that have not been detected in previous studies. Overall, this study contributes to the growing literature on CPR management by demonstrating the complex and context-dependent effects of strategic heterogeneity on cooperation.

4.6. LITERATURE CITED

Ambrus, A., & Pathak, P. A. (2011). Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics*, *95*(7), 500–512. https://doi.org/10.1016/j.jpubeco.2010.11.016

Amin, E., Abouelela, M., & Soliman, A. (2018). The Role of Heterogeneity and the Dynamics of Voluntary Contributions to Public Goods: An Experimental and Agent-Based Simulation Analysis. *Journal of Artificial Societies and Social Simulation*, *21*(1), 3.

Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *137*(4), 594–615. https://doi.org/10.1037/a0023489

Burlando, R. M., & Guala, F. (2005). Heterogeneous Agents in Public Goods Experiments. *Experimental Economics*, *8*(1), 35–54. https://doi.org/10.1007/s10683-005-0436-4

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*, *117*(3), 817–869. https://doi.org/10.1162/003355302760193904

Cheung, S. L. (2014). New insights into conditional cooperation and punishment from a strategy method experiment. *Experimental Economics*, *17*(1), 129–153. https://doi.org/10.1007/s10683-013-9360-1

Croson, R. T. A. (2007). Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games. *Economic Inquiry*, *45*(2), 199–216. https://doi.org/10.1111/j.1465-7295.2006.00006.x de Oliveira, A. C. M., Croson, R. T. A., & Eckel, C. (2015). One bad apple? Heterogeneity and information in public good provision. *Experimental Economics*, *18*(1), 116–135. https://doi.org/10.1007/s10683-014-9412-1

Fallucchi, F., Luccasen, R. A., & Turocy, T. L. (2019). Identifying discrete behavioural types: A re-analysis of public goods game contributions by hierarchical clustering. *Journal of the Economic Science Association*, *5*(2), 238–254. https://doi.org/10.1007/s40881-018-0060-7

Fehr, E., & Fischbacher, U. (2002). Why Social Preferences Matter – the Impact of non-Selfish Motives on Competition, Cooperation and Incentives. *The Economic Journal*, *112*(478), C1–C33. https://doi.org/10.1111/1468-0297.00027

Fehr, E., & Schmidt, K. M. (1999a). A Theory of Fairness, Competition, and Cooperation*. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Fehr, E., & Schmidt, K. M. (1999b). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

Fischbacher, U., & Gächter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review*, *100*(1), 541–556. https://doi.org/10.1257/aer.100.1.541

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404. https://doi.org/10.1016/S0165-1765(01)00394-9

Gächter, S., & Thöni, C. (2005). Social Learning and Voluntary Cooperation Among Like-Minded People. *Journal of the European Economic Association*, *3*(2/3), 303–314.

Gelcich, S., Guzman, R., Rodríguez-Sickert, C., Castilla, J. C., & Cárdenas, J. C. (2013). Exploring External Validity of Common Pool Resource Experiments: Insights from Artisanal Benthic Fisheries in Chile. *Ecology and Society*, *18*(3). https://www.jstor.org/stable/26269337

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Guido, A., Robbett, A., & Romaniuc, R. (2019). Group formation and cooperation in social dilemmas: A survey and meta-analytic evidence. *Journal of Economic Behavior & Organization*, *159*, 192–209. https://doi.org/10.1016/j.jebo.2019.02.009

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., & Sutter, M. (2008). Conditional cooperation on three continents. *Economics Letters*, *101*(3), 175–178. https://doi.org/10.1016/j.econlet.2008.07.015

Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences*, *102*(5), 1803–1807. https://doi.org/10.1073/pnas.0408759102

Ledyard, J. O. (1995). *Public Goods: A Survey of Experimental Research* (J. H. Kagel & A. E. Roth, Eds.; pp. 111–194). Princeton University Press. https://resolver.caltech.edu/CaltechAUTHORS:20171107-163442067

Lucas, P., de Oliveira, A. C. M., & Banuri, S. (2014). The Effects of Group Composition and Social Preference Heterogeneity in a Public Goods Game: An Agent-Based Simulation. *Journal of Artificial Societies and Social Simulation*, *17*(3), 5.

Marín, A., Gelcich, S., Castilla, J., & Berkes, F. (2012). Exploring Social Capital in Chile's Coastal Benthic Comanagement System Using a Network Approach. *Ecology and Society*, *17*(1). https://doi.org/10.5751/ES-04562-170113

Ones, U., & Putterman, L. (2007). The ecology of collective action: A public goods and sanctions experiment with controlled group formation. *Journal of Economic Behavior & Organization*, *62*(4), 495–521. https://doi.org/10.1016/j.jebo.2005.04.018

Ostrom, E. (1990, noviembre). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge Core; Cambridge University Press. https://doi.org/10.1017/CBO9780511807763

Rodriguez-Sickert, C., Guzmán, R. A., & Cárdenas, J. C. (2008). Institutions influence preferences: Evidence from a common pool resource experiment. *Journal of Economic Behavior & Organization*, *67*(1), 215–227. https://doi.org/10.1016/j.jebo.2007.06.004

Rustagi, D., Engel, S., & Kosfeld, M. (2010). Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. *Science*, *330*(6006), 961–965. https://doi.org/10.1126/science.1193649

Sobel, J. (2005). Interdependent Preferences and Reciprocity. *Journal of Economic Literature*, *43*(2), 392–436. https://doi.org/10.1257/0022051054661530

van Klingeren, F. (2022). Using player types to understand cooperative behaviour under economic and sociocultural heterogeneity in common-pool resources: Evidence from lab experiments and agent-based models. *PloS One*, *17*(5), e0268616. https://doi.org/10.1371/journal.pone.0268616

5. CONCLUSION

This study has demonstrated that strategic heterogeneity plays a fundamental role in shaping cooperation dynamics in CPR dilemmas. Through a combination of lab-in-the-field experiments and agent-based simulations, we provide insights into how the distribution of cooperative strategies within user groups influences the success or failure of co-management institutions. This approach offers a deeper understanding of the interactions between social norms, external enforcement, and individual strategic heterogeneity.

A key finding of this research is the capacity of high-performance CPR user groups to stabilize cooperation over time under conditions of non-deterrent external enforcement. This contrasts sharply with low-performance groups and non-unionized fishers, who exhibit an inability to maintain stable cooperation even with the introduction of enforcement. Analysis of individual strategies reveals that the proportion of UCs increases significantly under the enforcement treatment, particularly among high-performance groups, suggesting that external but non-deterrent enforcement can positively shift cooperative strategies towards more group-beneficial outcomes. However, the results from the first paper also suggest that these changes do not occur uniformly across groups, and that even in groups with increases in the proportion of more cooperative strategies, they may still fail to meet the challenge of sustaining cooperation over time.

Methodologically, this study contributes to the field by developing an agent-based model that simulates cooperation dynamics across different CPR contexts. By using experimental results to inform model parameters, we have been able to explore a broader range of possible outcomes and predict how variations in the prevalence of strategies may influence the stability of cooperation. These findings have important practical implications for co-management policy design, as they suggest that the effectiveness of such policies hinges not only on institutional enforcement but also on users' ability to adapt and shift their cooperative strategies in response to normative incentives.

Finally, this study underscores the importance of considering strategic heterogeneity in the design of management arrangements. Co-management policies that fail to account for the diversity of strategies within user groups may struggle to achieve sustainable levels of cooperation. In this sense, our results suggest that the most effective approaches to CPR governance should integrate enforcement mechanisms that, while imperfect, succeed in activating social norms and collective expectations of cooperation, thus facilitating a transition toward more stable and cooperative strategies over the long term. Future research should address this challenge, exploring how to strengthen the norm signaling of enforcement and which other mechanisms of imperfect (external or peer-to-peer) enforcement alter the strategic profiles of CPR user groups towards more cooperative and sustainable outcomes. Moreover, a critical and deeper challenge that must also be explored is which organizational or institutional factors shape the different responses of CPR user groups to changes in resource management incentives, and if strategies are also sensitive to resource scarcity or just to social behavior.