

Payments or Persuasion: Common Pool Resource Management with Price and Non-price Measures

Jason Delaney¹ · Sarah Jacobson²

Accepted: 27 April 2015 / Published online: 16 May 2015
© Springer Science+Business Media Dordrecht 2015

Abstract We use lab experiments to study policies that address common pool resource overuse. We look at a price mechanism, specifically a Pigouvian subsidy, and four non-price interventions. The non-price policies are information alone, information with a normative message, communication alone, and normative messages with communication allowed. In all experiment sessions, no intervention occurs in the first seven and last seven rounds, allowing us to examine the effects of introducing and taking away a policy. The subsidy leads to near-efficient extraction, but surprisingly leads groups that were not over-extracting to also reduce extraction. This over-compliance decreases efficiency, although on net the subsidy is the most efficiency-enhancing intervention. Information provision, communication, normative appeals, and normative appeals combined with communication all reduce over-extraction (though by less than the subsidy) without exacerbating over-compliance; however, the effects of information alone and communication alone are small and not robust. The non-price policies cause a decline in over-extraction of from 0.549 (information) to 11.441 % (normative appeals with communication). These effects are of the same order of magnitude as the effects seen in major field studies of conservation messaging. The subsidy has the worst persistence properties (after the intervention ceases), while normative messages with communication have the best.

Electronic supplementary material The online version of this article (doi:[10.1007/s10640-015-9923-z](https://doi.org/10.1007/s10640-015-9923-z)) contains supplementary material, which is available to authorized users.

We thank James Cox, Cary Deck, Ragan Petrie, Vjollca Sadiraj, Christian Vossler, two anonymous referees, and colleagues and seminar participants for comments and suggestions. We thank Georgia State University and Williams College for funding.

✉ Sarah Jacobson
sarah.a.jacobson@williams.edu

Jason Delaney
jdelaney@ggc.edu

¹ School of Business, Georgia Gwinnett College, 1000 University Center Lane, Lawrenceville, GA 30043, USA

² Department of Economics, Williams College, 24 Hopkins Hall Dr., Williamstown, MA 01267, USA

Keywords Common pool resource · Communication · Information · Pigouvian subsidy · Social norms · Voluntary cooperation · Laboratory experiment

JEL Classification H41 · H21 · C91 · C92 · D62 · D83

1 Introduction

Environmental and natural resource settings are fraught with common pool resource (CPR) problems, whether the focus is on local artisanal fisheries or on the global challenge of climate change. These problems occur when a good is rival but not excludable. Under such conditions, the marginal cost of consumption incurred by the decision maker differs from the marginal social cost borne by the group at large. This can lead to congestion and overuse, and, in some cases, the destruction of the resource.¹ The Pigouvian (Pigou 1920) solution to externalities like these can theoretically achieve efficiency. Unfortunately, price-based solutions (like Pigouvian taxes and subsidies) to CPR overuse are difficult to implement, so policy has often turned to non-price interventions. Information provision and appeals to social norms can nudge people to voluntarily reduce CPR extraction.² However, the literature has not yet provided a rigorous comparison of the performance of these non-price policies to a price-based intervention in a CPR setting.

We use lab experiments to fill this gap. Our contributions are as follows. First, we provide the first “apples to apples” comparison between a theoretically-efficient price intervention (a Pigouvian subsidy) and several non-price interventions (information provision, communication, normative messaging, and normative messaging with communication) in a CPR setting. Second, we take a careful look at efficiency. To do this, we consider what effects the treatments have on groups that are over-extractors, optimal-extractors, and under-extractors, since average extraction across groups need not be informative about efficiency. We also compare persistence of each policy’s effects after the policy has been discontinued. As we discuss below, elements of these questions have been answered in existing literature, but such studies have been in less controlled settings or they have not allowed for comparison across the different treatment types.

We find that, while all treatments reduce overuse, price has the largest effect. The normative message with communication has the largest effect of any of the non-price interventions. Though it is the most effective, the price intervention has drawbacks. It treats everyone equally, and as a result groups that initially underuse the resource reduce use further, which is inefficient. The subsidy also has the worst persistence properties, while the normative message with communication has the best. Finally, while the subsidy has the strongest efficiency effects, the non-price interventions successfully target overuse, also significantly increase efficiency, and are less likely to suffer efficiency losses through monitoring, enforcement, and administration costs. These results are important because they provide insight into the strengths and weakness of crucial tools in the environmental policy toolbox.

¹ “Congestion” and “overuse” are used somewhat interchangeably in the literature on CPRs. In this paper, we refer primarily to “overuse” as the CPR problem we address results from excessive consumption within a group of fixed size, rather than “congestion,” which could mean a superabundance of users.

² Another tool for fighting externalities is decentralized Coasian bargaining (considered experimentally in, e.g. Harrison et al. 1987). Various work by Ostrom and her coauthors has shown that varying other parameters (e.g. endowment) can lead to changes in appropriation levels, all else equal; so while we focus here on price and information, other elements can affect behavior.

CPR problems and the policies that address them are complex and multifaceted. As a consequence, lab experiments are an essential tool for studying human behavior and institutions in these settings. Much work in this area follows pioneering studies by Elinor Ostrom and coauthors (e.g. in [Ostrom et al. 1994](#)). We follow in this tradition. Performance in CPR settings without policy intervention has been generally shown to be worse than the social optimum but better than the Nash equilibrium,³ as found by [Ostrom et al. \(1994\)](#), [Budescu et al. \(1995\)](#), and [Janssen et al. \(2011\)](#), for example, although [Walker et al. \(1990\)](#) find that the CPR is dissipated by more than the Nash prediction. Ostrom's work also shows that CPRs are often successfully and sustainably governed without financial incentives ([1992](#); [2000](#); [2002](#)).⁴

We must understand why non-Nash behavior occurs in these settings so that policy can account for it. Other-regarding preferences such as altruism ([Dawes 1980](#)), conditional cooperation ([Gächter 2007](#)), inequity aversion ([Falk et al. 2002](#)), or even spite ([Casari and Plott 2003](#)) may drive this non-Nash behavior. Social norm conformance has been shown to be important, as discussed below. In other cases, as in [Bru et al. \(2003\)](#), it's not entirely clear why people behave as they do, although confusion may be a factor. If people respond to non-monetary incentives, non-price instruments may reduce congestion and overuse, and Pigouvian instruments based solely on pecuniary costs and benefits may be inefficient. [Johansson \(1997\)](#) discusses how Pigouvian taxes must be modified under altruism, for example. In our results, we not only replicate the result from the literature that many subjects over-extract less than (i.e. "cooperate" more than) the Nash prediction, but also find that a nontrivial population under-extracts, as others have found. Policies like the Pigouvian subsidy ignore these non-Nash behaviors, and the result is that inefficient under-extraction is exacerbated in our context.

Policymakers have used a variety of policies to address CPR overuse, including both price and non-price interventions. These policies have been studied independently but never before directly compared.

Pricing externalities is considered a first-best solution (although this may ignore social preferences), and prices have been implemented in policy as taxes (e.g. British Columbia's carbon tax) and subsidies (e.g. the United States Conservation Reserve Program and Renewable Electricity Production Credit). Lab studies have found that Pigouvian instruments increase efficiency to nearly optimal levels when externalities exist ([Plott 1983](#); [Cochard et al. 2005](#)). On the other hand, [Anderson et al. \(2007\)](#) find persistent inefficiency with congestion pricing in an entry and exit game, although this is due to variance in entry. In a non-point source setting (where the regulator can only monitor group, not individual, behavior), [Vossler et al. \(2006\)](#) and [Cochard et al. \(2005\)](#) find over-compliance in response to fines because of collusion. However, price-based instruments, particularly taxes, can be politically difficult to implement, in part because of tax aversion (e.g. [Kallbekken et al. 2011](#); [Heres et al. 2013](#)). More fundamentally, price tools are often difficult to implement because they require a level of monitoring and enforcement that is infeasible for many CPRs.

Non-price policies have an important role because they can both leverage non-price drivers of behavior and potentially avoid some of the drawbacks of price-based instruments. Information provision programs, such as Spare the Air ([Cutter and Neidell 2009](#)) and Energy Star,

³ Here and elsewhere, we use "Nash equilibrium" to refer to the Nash equilibrium under the standard assumption that agents derive utility only from their own payoff.

⁴ [Suter et al. \(2012\)](#) show that an accurate model of groundwater as a CPR involves less externality and more private cost burden than the standard model, and in a lab experiment they show that this model (as one would expect) yields less over-consumption.

tell people how their actions affect others in hopes that this will induce conservation.⁵ Haynie et al. (2009) find that information enables cooperation in an Alaskan fishery, although information's effect declines over time. Some studies have tried to use observational data to study the impact of informational programs in comparison to the effect of prices, such as Renwick and Green (2000) who investigate of water use in California, Dulleck and Kaufmann (2004) who study Irish residential electricity, and Coleman (2009), who studies water use in Salt Lake City, but identification in these cases is difficult because treatment is not exogenous.

Some recent programs have used appeals to social norms, with a prominent example being the Opower energy conservation campaign. Through Opower, residential electricity customers get, with their electric bills, messages that perform social comparisons and award "smiley faces" or "frowny faces" based on a household's energy use. These policies build on the findings, e.g. in Schultz et al. (2007), that social norm conformance can drive good behavior in CPR settings. Rodriguez-Sickert et al. (2008) find that even a fine that is not implemented affects behavior, perhaps because a norm is established when subjects consider the fine. Osés-Eraso and Viladrich-Grau (2011) find that subjects reward each other for pro-social behavior in a CPR, and rewards, like punishments (Ostrom et al. 1992), are a form of norm enforcement. Game structure and context may determine which norms are viewed as appropriate: Cox et al. (2009) find that framing a resource as commonly (rather than privately) owned increases efficiency.

These results extend beyond the laboratory. In artefactual field experiments, Velez et al. (2009) show that subjects balance self-interest with conformity, while Ferraro and Price (2013) show that people reduce water consumption when treated with normative messages containing social comparisons. LaRiviere et al. (2014) compare a price to a normative message with social comparisons in promoting home energy audits, and estimate that a social nudge is as effective as a \$40 rebate. Ayres et al. (2013) and other studies find that Opower's use of social norm-laden messaging significantly reduces energy usage. In another study of Opower, Allcott and Rogers (2014) show that a normative intervention has a persistent impact on household energy usage even after the intervention is discontinued, and Ferraro et al. (2011) also show a persistent effect of normative messaging over a 2-year period in comparison to no persistence for messages with simple information or appeals to prosocial preferences. Our study echoes these results, in that we find that non-price interventions do have pro-social effects and generally have better persistence properties than the price intervention. Our results differ from the literature in that, although we follow the literature's guidance on effective normative messages, our normative messaging treatment is barely (and generally not significantly) more effective than simple information—but when communication is allowed, normative messages are significantly more effective.

Nonpecuniary interventions have thus been shown to have small but significant effects. These interventions can also be inexpensive to implement: they do not necessarily require monitoring and enforcement at the individual level because they can involve merely blanket-ing an entire group with the same message or information. Our contribution to the literature is to compare messaging policies in a CPR setting against two benchmarks: no treatment and Pigouvian subsidy. Such a comparison can only be done with real precision in the lab. Further, the lab allows us to study several subtle but important features of the policies' effects. We study the policies' effects on average resource appropriation, but we also look at under-appropriation as well as over-appropriation, the effects of the policies on different types of

⁵ We are most interested in use of information to encourage cooperation, but information can also be used to enable coordination, as in Anderson et al. (2007), who find that information on current congestion levels can help subjects coordinate to reduce variance and therefore increase efficiency.

groups, persistence of gains after the intervention ceases, and economic efficiency before, during, and after interventions. Our results highlight the roles of conditional cooperation and norm conformance in explaining behavior, and show that heterogeneity in behavior has important implications.

The remainder of the paper proceeds as follows. Section 2 lays out a model of the CPR game of interest, providing a theoretical motivation for each of the policy interventions. Section 3 describes the experiment design. Section 4 explores the results empirically, and Sect. 5 concludes.

2 Theory

We study behavior relating to CPRs. A CPR is rival, in that one agent’s consumption reduces the stock available to others, and non-excludable, in that it is costly to prevent consumption. A CPR’s rivalness creates a negative consumption externality (an overuse externality, as in Brown 1974). As a result, the privately optimal level of consumption differs from the social optimum, and inefficiency persists in equilibrium.

2.1 A Common Pool Resource Appropriation Game

A group shares a CPR. In each period, each person in the group must decide how much she wants to extract from the resource. Each agent $i \in (1, \dots, n)$ has an endowment of z_i tokens, which represents a limited resource that can be used for CPR extraction. In each round, each agent chooses an amount $x_i \in (0, \dots, z_i)$ to allocate toward CPR extraction. The total group extraction from the CPR is Σx_i . For each token she spends on the CPR, agent i forgoes constant opportunity cost α . Her return from spending x_i tokens on the CPR depends on her token expenditure and on the total spent on the CPR by all agents: $g(x_i, \Sigma x_j)$. The payoff to agent i is the return from tokens not spent on the CPR plus the return from CPR appropriation.

To model an overuse externality, we define the payoff to the CPR as $g(x_i, \Sigma x_j) = (\beta - \gamma \Sigma x_j)x_i$. Agent i ’s per-token return from the CPR starts at $\beta > 0$ and decreases by $\gamma > 0$ for each token she or any other group member has spent on the CPR. A rational self-interested agent chooses x_i to maximize her payoff:

$$\pi(x_i, \Sigma x_j) = \alpha(z_i - x_i) + (\beta - \gamma \Sigma x_j)x_i$$

This is mathematically identical to the payoff function used in Ostrom et al. (1994) and all the underlying incentives are identical, although our framing is more like that of the traditional linear voluntary contributions mechanism public goods game (Marwell and Ames 1979). The nonlinearity of the CPR payoff function generally yields an interior solution ($0 < x_i < z_i$).

This game has no dominant strategy. There exists both a symmetric Nash equilibrium ($x_i^{NE} = \frac{\beta - \alpha}{(n+1)\gamma}$) and a symmetric social optimum ($x_i^{SO} = \frac{\beta - \alpha}{2n\gamma}$).⁶ The socially optimal level of consumption and the Nash equilibrium level of consumption are only identical for $n = 1$ (in which case there is no externality) or $\beta \leq \alpha$ (in which case even the uncongested CPR is not worth the opportunity cost that must be incurred to exploit it). We will henceforward only consider groups larger than one, under conditions with an incentive to appropriate a non-zero amount from the CPR. The optimal (Pigouvian) subsidy is $\delta = (\beta - \alpha) \frac{n-1}{2n}$. At the optimum, the total social return from the CPR is $G^{SO} = \sum_{i=1}^n g(x_i, \Sigma x_j) = \frac{(\beta^2 - \alpha^2)}{4\gamma}$.

⁶ We derive these, the optimal (Pigouvian) subsidy, and the total resource yield at the optimum in the online supplementary material.

We define the *gross efficiency* of any level of resource yield $G(\cdot)$ as its percentage of the social optimum: $\frac{G(\cdot)}{G^{SO}}$. An alternative definition of efficiency is the efficiency of the combined system, including both CPR use and consumption of the outside option. This can be considered the efficiency of the resource use net of the opportunity cost, or *net efficiency*, and is the same as considering overall group payoffs.⁷ Net efficiency has the advantage of capturing a more expansive definition of economic costs and benefits, but has the disadvantage of being sensitive to the size of the CPR relative to the size of the economy—if the CPR is relatively small, it can have little effect on overall efficiency. In our experiment, however, we normalize the return to the outside option by setting it to zero. Thus, in our later results and discussion, the two measures of efficiency are equal, and so in the following, we simply refer to “efficiency.”

2.2 The Search for Efficiency

We expect private consumption to be above the social optimum. Pigou (1920) showed that a subsidy can induce the socially optimal choice, and that the optimal subsidy is the difference between the marginal private cost and the marginal social cost at the optimal quantity. To introduce a subsidy, we add a fixed per-token amount (δ) to the return to the outside option.⁸ The transformed payoff function is:

$$\pi(x_i, \Sigma x_j) = (\alpha + \delta)(z_i - x_i) + (\beta - \gamma \Sigma x_j)x_i$$

The implementation of a Pigouvian subsidy runs into two problems. First, funds must be raised to provide the subsidy, and most revenue-raising taxes create inefficiency. Second, CPR users must be identified and their consumption monitored. However, one of the key characteristics of a CPR that makes it non-excludable is the possibility of hidden action (Arrow 1984). Monitoring and enforcement may be impractical or excessively costly for CPRs of interest.

In place of a Pigouvian instrument, a policy that presents information about the CPR’s overuse to the CPR users could reduce overuse by engaging naturally existing social preferences, including altruism or efficiency-seeking. Experimental evidence suggests that people respond to not only their own private costs but also the costs borne by others (Goeree et al. 2002). Such policies have advantages over a tax or a subsidy: political economy presents fewer problems, monitoring of individual behavior is no longer necessary, and they are less complex (and costly) to implement.

Direct appeals to social norms may be even more effective than simple information provision. Normative appeals have been successful in field experiments (Ayres et al. 2013; Ferraro and Price 2013). Evidence suggests that the most successful normative appeals combine positive descriptive messages (“many do choose socially desirable behavior”) with negative injunctive messages (“do not engage in socially undesirable behavior”) (Schultz et al. 2007). These messages engage social preferences as information presentation does, but also bring to bear social comparisons (Festinger 1954).

⁷ In this context, recalling that x^{SO} is the level of CPR consumption at the social optimum, net efficiency would be $\frac{G(\cdot) + \sum_{i=1}^n \alpha(z_i - x_i)}{G^{SO} + \sum_{i=1}^n \alpha(z_i - x^{SO})}$.

⁸ When the subsidy is added, the Nash equilibrium of the modified game shifts to the socially optimal condition of the original game, and as an externality still exists the modified game has what would appear to be a new social optimum. The goal of policy is to reach the social optimum of the un-modified game. At the same time, the subsidy is not assumed to inherently have any efficiency costs or gains, but is assumed to be a transfer. The efficiency costs of raising funds to pay the subsidy are discussed in Sect. 4.3.

2.3 Behavioral Motives

If people play differently than the Nash prediction, what motives will drive behavior? Most of these motives relate to theories of other regarding preferences, as we discuss below. However, risk preferences also play a role. Since allocating tokens toward the CPR carries a strategic risk while the outside option yields a certain return, risk averse people may appropriate less from the resource as compared to the Nash prediction. Any intervention that reduces strategic risk should cause risk averse people to appropriate more, while any intervention that increases strategic risk should cause them to appropriate less.

Ostrom (1998) notes that the rational model of choice in social dilemmas should be extended to include a variety of social motives that have been studied in theory and experiment across disciplines. In our setting, relevant social motives include the following. First, pure or impure altruism (Andreoni 1989) could cause agents to want to increase the payoffs of their group members—either because they derive utility directly from other players' higher payoffs, in the case of pure altruism, or because they derive utility from the act of contributing to others, in the case of impure altruism. Second, conditional cooperation (Gächter 2007) is a general form of reciprocal behavior that could cause lower appropriation than the Nash at first, and then an increase in appropriation as people notice others over-appropriating. More general phenomena of direct reciprocity (Cox et al. 2007) and indirect reciprocity (Nowak and Sigmund 1998; Nowak and Sigmund 2005) could drive a similar pattern. Finally, agents may be averse to violations of social norms (Schultz et al. 2007), and thus if norms of cooperative behavior are made salient then over-appropriation may drop.

3 Experiment

In the experiment, subjects interact in three-person groups ($n = 3$) for 21 rounds. The composition of the groups is fixed for the entire session (there is no rematching). In each round, each subject gets an endowment of $z = 10$ tokens and must decide how to allocate those tokens between two accounts. The RED investment represents the outside option and pays a fixed per-token amount to just the subject herself. Each token put in the BLUE investment, which represents the CPR, pays a per-token amount that depends on the total number of tokens the group invested in BLUE.

Each subject's payment for the session is the sum of the payoffs she earns in each round. The session consists of three seven-round phases: a first baseline, a treatment phase, and a second baseline. Subjects know that the game will last for 21 rounds and can observe when a change in treatment conditions has occurred (either by having been told that the payoffs have changed, or by receiving a message in the nonpecuniary treatments), but they are not otherwise told when a treatment will begin or end or how it will differ from preceding rounds (to reduce experimenter demand effects). The second baseline lets us observe trends in behavior such as learning or the decay in cooperation that has been widely documented in public goods games (e.g. Gächter 2007). It also lets us consider persistence after an intervention has been removed.

There are five treatments, each described in detail below. The design is across-subject: each subject receives just one treatment for all seven periods of the treatment phase. In all treatments, the per-token baseline RED payoff is $\alpha = \$0.00$,⁹ the per-token starting BLUE

⁹ A return of zero for the outside option is unusual in the literature. However, increasing the return to the outside option would require an offsetting increase in the return to the CPR, as the difference in these returns

payoff is $\beta = \$0.36$, and the per-token BLUE externality parameter is $\gamma = \$0.01$. In the baseline phases, these parameters give a symmetric Nash solution of $x_i = 9$ tokens, while the social optimum is $x_i = 6$. The social optimum and Nash are interior to the choice space and are far enough from each other to allow for statistical inference. Because of the nonlinear payoff structure, the differences may not appear large to those more familiar with linear public goods games: at the social optimum, subjects earn \$26.88 each, but subject earnings at the Nash equilibrium outcome are \$22.26. The range of feasible payoffs is much wider: \$0.00–\$54.60, or \$4.20–\$54.60 for the subsidy treatment. As we discuss later, our results indicate that these incentives are strong enough to induce meaningful behavioral responses.

3.1 Subsidy Treatment

In the subsidy treatment, the return to the outside option is augmented with a per-token RED subsidy of $\delta = \$0.12$. This is the Pigouvian subsidy: it is the difference between the private and social marginal costs at the social optimum. The subsidy is not framed explicitly as a subsidy, but is simply a different return to RED. We wish to avoid any framing effects that could come about from describing this as a subsidy so that the behavioral response we observe is only a response to a change in prices. With this subsidy, the symmetric Nash equilibrium is $x_i = 6$, which yields the socially optimal outcome.

3.2 Information Treatment

In the information treatment, subjects receive a message that depends on the group level of CPR appropriation in the previous period. If over-consumption occurred, subjects see:

"During the last period, your group earned a total of $[\sum \pi_{j,t-1}]$. The maximum your group could have earned was \$3.24. Your group earned LESS than it could have in that period. To increase your group's total payoff, your group should REDUCE its investment in the BLUE investment."

If under-appropriation occurred, the word "REDUCE" is replaced by the word "INCREASE." If optimal appropriation occurred, subjects receive the message:

"Your group earned as much as it could have in that period. Your group is investing in the BLUE investment at the right level to maximize your group payoff."

3.3 Normative Messaging Treatment

In the normative messaging treatment, groups that were over-appropriating in the previous round receive the following message:

Footnote 9 continued

governs strategic considerations. This would amount to a strategically irrelevant transfer from experimenters to subjects. Setting the opportunity cost to zero thus allows us to make the payoff space as responsive to different levels of investment in the CPR as possible. In addition, this removes any ambiguity about the definition of efficiency as discussed above.

"Your group earned LESS than it could have in last period! :(In previous experiments, many groups contributed LESS to BLUE than your group did, and the members of those groups earned more as a result. To boost your group's earnings, DO NOT put so much in BLUE!"¹⁰

Groups that were appropriating less than the social optimum receive a similar message with "RED" replacing "BLUE." Groups with optimal appropriation receive the message:

"Your group earned as much as it could have last period! :) To maintain your group's earnings, do not change the total number of tokens in the BLUE investment!"

The information treatment message is not wholly without normative content, since it uses the word "should," although the "should" is conditional and the message does not explicitly endorse the end that this guidance would achieve. However, the normative treatment uses a stronger norm-laden message and follows the lessons from the literature of using positive descriptive and negative injunctive language. This should cause the normative treatment to affect behavior more than the information treatment.

On the other hand, while the normative message is not without information content, the information treatment provides more specific and detailed quantitative information about the extent of group over- or under-appropriation as compared to the normative messaging treatment. As a result of this, the information message received by a group will change from round to round as group appropriation changes while the normative message only changes when a group makes a major transition like a change from overextracting to optimal-extracting. It is unclear what effects these differences will have on behavior.

This normative message is essentially a social signal. We deliberately implement this message in the context of a laboratory experiment in which we can exclude reputation and social pressure, which might act in a non-lab setting to counteract or, more likely, reinforce the effect of the message. The effect that we measure, therefore, is a pure effect of the invocation of a social norm.

3.4 Communication Treatment

In the communication treatment, each subject can send a message to everyone in her group after each round. To avoid risks of subjects' compromising their anonymity or promising post-experiment reward or retribution, we only allow subjects to choose from a list of pre-defined messages, and then randomly re-order those messages before transmitting them from the group. This parallels the other nonpecuniary treatments, in that subjects cannot identify other members of the group, and that the messages are sent after observing previous group-level behavior.¹¹ The messages (listed in Table 3) provide all sensible combinations of instructions for contribution behavior (increase, decrease, or stay the same) with three different attitudes (neutral, friendly, or angry), as well as messages expressing uncertainty

¹⁰ This statement was true for all of our subjects because we ran the normative treatment after having run some sessions of the other treatments.

¹¹ Face-to-face communication would not meet these criteria, for example, and would risk a loss of experimental control. In this treatment, for a sensible tradeoff between flexibility and control, we observe all subject behavior and use a decentralized communication institution to compare to the centralized messages sent in the other nonpecuniary treatments.

and apology, and the option to send no message. Everyone in the group receives a randomly-ordered list of the messages sent by all three group members (including herself) before the next round starts. Subjects cannot link messages to appropriation behavior of any individual group member.

3.5 Normative Messaging with Communication Treatment

To study the effect of normative messaging when subjects are embedded in a more social context, we send normative messages based on group extraction levels as in the normative messaging treatment, but subjects can also send messages to their group members as in the communication treatment. We refer to this treatment briefly as normative + communication.

3.6 Experiment Execution

Subjects are recruited to a session. (Each session runs only one treatment, so across-treatment analyses compare data from different sessions.) Upon arrival, they are signed into the laboratory. The experiment is computerized (programmed in Visual Basic by the authors), so subjects are seated at computers. Subjects read instructions on their screens while the experimenter reads them aloud.¹² The instructions describe the game structure and the computer interface. Subjects interact with a software tutorial and then practice using the software for as long as they like.¹³ In the practice rounds, the computer deterministically plays as the “rest of the group” (putting 0 tokens in the BLUE investment in the first round, 1 in the second round, continuing up through 20, then restarting at 0 tokens).

After all subjects have stopped practicing, subjects are randomly assigned into three-person groups in which they remain throughout the experiment. Subjects know that they will not learn the identities of their group members.¹⁴ They make the series of twenty-one “investment” decisions as described above: seven rounds of baseline, seven rounds of a treatment (one of the five described above), and seven rounds of a second baseline. During each round, the decision is made easier by the availability of a calculator that allows each subject to see what their earnings would be given any combination of their chosen appropriation and a guess about the decisions of the other subjects. After each round, a history table on the screen adds information about the preceding round, and this information stays in the table for the remainder of the experiment. For each round, the table shows what this subject chose, the total number of tokens put in BLUE by the rest of the group, this subject’s earnings for the rounds, and this subject’s cumulative earnings.¹⁵

The sessions are run with a double-anonymous protocol to reduce experimenter influence (see, for example, [Cox and Deck 2005](#)): a subject learns what decisions her group members made but never knows which other people in the room were in her group, and the experimenter can never link this subject’s decisions to her identity. After all decisions, subjects complete a questionnaire and then receive payment alone and anonymously.

¹² Instructions are available on the corresponding author’s website and in the online supplementary material.

¹³ In the instructions and the tutorial, the returns to RED and BLUE differ from those used in decision-making rounds to discourage subjects from anchoring to any choices they see as examples.

¹⁴ The instructions imply but do not explicitly state that groups will remain fixed for all rounds. Similarly, subjects are not explicitly told that payoff functions are identical for all group members, although it is implied. Subjects seemed to understand both of these features.

¹⁵ There was no explicit test of subject understanding; however, subjects did not appear confused, and the data that we will present—such as the fact that average contributions drop to the new Nash when a subsidy is imposed and rise again to the new Nash when the subsidy is removed—indicate general understanding.

4 Results

The experiment was conducted in ten sessions with 237 subjects in total at Georgia State University's Experimental Economics Center (ExCEN) in April 2009, July 2011, and October 2014. Each session used 21–27 subjects and lasted about an hour and a half. Individual earnings, including a \$5 show-up payment, ranged from \$17.98 to \$40.60 and averaged \$25.35. Most subjects used the practice rounds: only 61 subjects (25.738 %) took zero rounds of practice, and the average number of practice rounds was 7.059.

Because subjects within a group may influence each other, and because we are primarily interested in the efficiency of total CPR use, all analysis we perform is at the group, rather than individual, level. Total group appropriation from the CPR (contribution to the BLUE investment) is the sole determinant of efficiency in the sense of maximizing resource yield. In other words, individual variation in appropriation within a group does not affect resource yield and therefore efficiency.¹⁶

On the other hand, while variation in appropriation within a group has no efficiency impact, variation in appropriation across groups can affect efficiency. Since this game (like most CPR games) has a nonlinear overuse cost function, an improvement in average extraction may not imply an increase in efficiency. For example, if average extraction moves closer to the optimum but there is an increase in the average (absolute value of the) distance between group extraction levels and the optimal, this will cause efficiency to decrease. This may happen if the average improves as a result of reductions by groups that were already under-extracting. Therefore, while we start by examining group appropriation averaged across groups, these averages may be misleading in terms of actual efficiency. We follow by studying responses to treatment for groups that are over-extracting, optimally extracting, and under-extracting, and then finally we directly measure efficiency (resource yield).

Figure 1 shows group appropriation decisions for the three treatments as the rounds of the experiment progress. Recall that rounds 1–7 are the first baseline, 8–14 are the treatment, and 15–21 are the second baseline. Since the endowment is $z = 10$, the maximum possible appropriation in a round is 30. Horizontal lines on the plots in Fig. 1 show the social optimum (18 tokens) and the Nash equilibrium (27 tokens). Panels (a) through (e) show the distribution of group appropriation in each period for each treatment, while panel (f) plots the period averages of group appropriation decisions for all five treatments together.

Over-appropriation is clearly a problem in these lab CPR's, but (at least in the first two phases of the experiment) we observe a substantial number of groups that under-appropriate as well.¹⁷ For each treatment, a treatment effect is visible. The subsidy seems to cause appropriation to drop to the social optimal (the new Nash), and after the subsidy is removed appropriation rises all the way to the Nash level. By comparison, the non-price interventions look dramatically less effective than the subsidy, although they do seem to reduce appropriation. In each case, cooperation also seems to decay as the rounds progress, as all lines slope upward (except for the subsidy after the first baseline, in which case there essentially is no voluntary cooperation). This cooperative decay may be caused by conditional cooperation (Gächter 2007).

¹⁶ Results are robust to individual-level analysis.

¹⁷ Most decisions are not corner choices: only 18 (0.362 %) individual-round choices are zero, only 1,355 of 4,977 (27.225 %) individual-round choices are full endowment (10 tokens), and 32 of 1,659 (1.929 %) group-round aggregate choices are group endowment (30). The rate of choosing full endowment appropriation varies directly with the average extraction across treatments and phases. As mentioned below, regression results hold for a Tobit specification that accounts for censoring, so censoring is not driving our results.

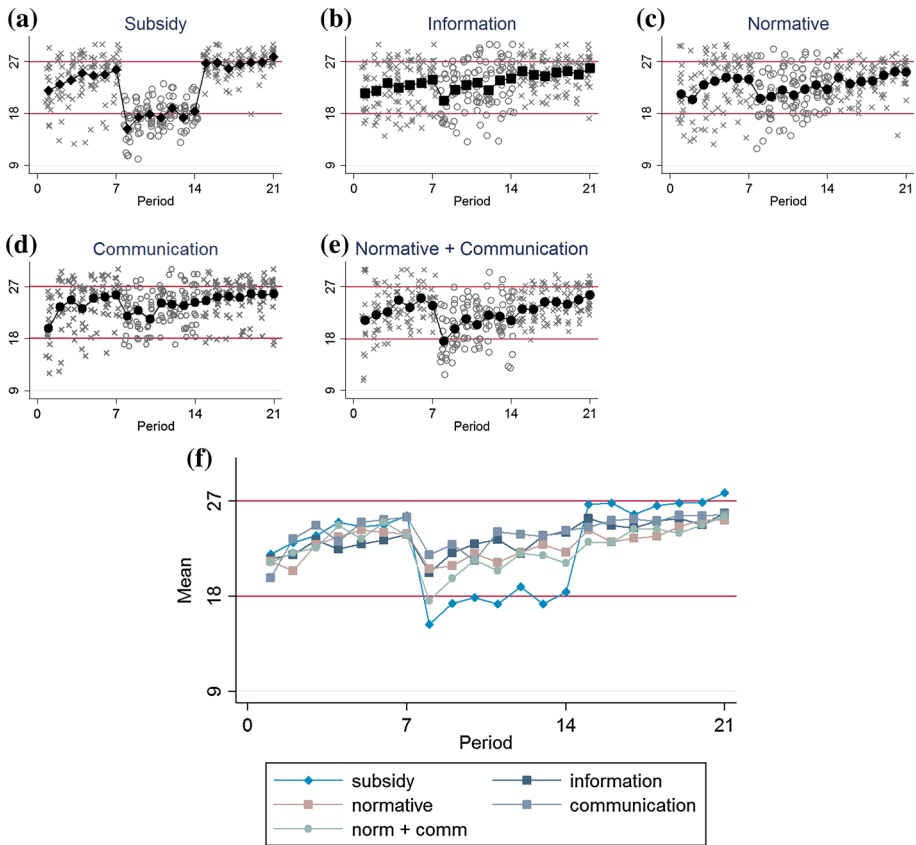


Fig. 1 Group appropriation by round (in tokens). **a–e** Group appropriation averages and observations for the subsidy, information, normative message, communication, and normative + communication treatments respectively. **f** Mean group appropriation for all treatments

4.1 Descriptive Results and Comparisons

Table 1 summarizes group-level appropriation averaged across all rounds of each phase of each session type, and Table 2 provides *p* values of additional associated tests. These tests are not only performed at the group, rather than individual, level, but they are also performed on data aggregated across rounds so they are robust to and abstract away from across-round dynamics; they are also non-parametric so they do not depend on functional form.¹⁸

First, we note the evidence of a decay of cooperation across the treatments. In the first baseline, average group appropriation across all treatments is 23.389 tokens (out of a maximum of 30). This drops during the treatment periods and then rises back to a mean of 25.128 tokens during the second baseline. This decay is statistically significant for the pooled data and for all treatments except normative messaging with communication (while normative messaging alone shows a borderline significant decay).

Next, we examine treatment effects. CPR appropriation in the first baseline is basically statistically identical across all treatments (though the difference between the two treatments

¹⁸ The results are robust to a panel regression specification.

Table 1 Group appropriation levels

	Pooled	Subsidy	Information	Normative	Communication	Norm. + comm.
Baseline 1	23.389 (2.701)	24.116 (2.933)	22.777 (2.779)	22.933 (2.946)	23.83 (2.799)	23.259 (2.066)
Treatment	21.125 (3.082)	17.518 (1.404)	22.652 (3.044)	21.733 (2.543)	23.161 (2.359)	20.598 (2.258)
Baseline 2	25.128 (2.165)	26.759 (1.030)	25.098 (2.099)	24.238 (2.111)	25.313 (2.605)	24.179 (1.835)
Treatment effect versus base. 1		-6.598	-0.125	-1.200	-0.669	-2.661
<i>N</i> (groups)	79	16	16	15	16	16
Within-treatment test <i>p</i> values						
Base. 1 = Treatment	<0.001	<0.001	1.000	0.094	0.301	0.007
Treatment = Base. 2	<0.001	<0.001	<0.001	0.002	<0.001	<0.001
Base. 1 = Base. 2	<0.001	0.001	<0.001	0.094	0.021	0.266
Base. 1 = Nash	<0.001	<0.001	<0.001	0.001	<0.001	<0.001
Treatment = Nash		0.088	<0.001	<0.001	<0.001	<0.001
Base. 2 = Nash	<0.001	0.363	0.003	<0.001	0.012	<0.001

Standard deviations in parentheses. Values in number of tokens in total group appropriation (0–30). Nash is 27 tokens except in the treatment phase of the subsidy when it is 18 tokens. Within-treatment *p* values are *p* values of Wilcoxon signed-rank tests. All tests use group-level data

with communication is barely marginally significant), as expected, and is significantly less than the Nash prediction (i.e. there is voluntary cooperation) on average.¹⁹ In all treatments, appropriation is significantly greater in the baseline (pooled across first and second baselines) as compared to the treatment phase (*p* < 0.05 in all cases).

Our benchmark price treatment, the subsidy, is extremely effective. The subsidy gives appropriation that is much lower than any other intervention. In fact, the subsidy-induced appropriation is slightly but significantly below the social optimum, which is the Nash equilibrium for this treatment.²⁰ After the subsidy is removed, however, appropriation rises to the highest level we observe in the study, a level indistinguishable from the Nash prediction.

By comparison, the non-price interventions have small but varying effects. The information and communication treatments are the least effective; normative messaging appears more effective; and the normative + communication treatment is the most effective (although appropriation is not significantly less than in the normative messaging treatment).

Given that the literature (e.g. Goldstein et al. 2008; Ferraro and Price 2013) has found significant effects from adding normative language to simple information, it is remarkable that the normative treatment did not have significantly greater effects than the information treatment, even though the former incorporated “best practices” in normative messaging. This result is at odds with the literature, although the difference between the treatments might be

¹⁹ We compare average behavior to the Nash prediction simply as a benchmark. Individual behavior may diverge from treatment averages so an average of Nash behavior would not necessarily imply Nash play.

²⁰ This under-extraction could theoretically occur because risk-averse agents who are uncertain about how their group members will play may shift their allocation from the CPR to the outside option, since the latter has a certain return.

Table 2 Tests of appropriation levels across treatments

	Information	Normative	Communication	Norm. + comm.	<i>N</i> (groups)
Subsidy					
Base. 1	0.117	0.185	0.692	0.101	16
Trtmt	<0.001	<0.001	<0.001	<0.001	16
Base. 2	0.010	<0.001	0.076	<0.001	16
Info.					
Base. 1		0.906	0.235	0.749	16
Trtmt		0.286	0.792	0.022	16
Base. 2		0.277	0.521	0.152	16
Norm.					
Base. 1			0.259	0.737	15
Trtmt			0.063	0.363	15
Base. 2			0.082	0.782	15
Comm.					
Base. 1				0.097	16
Trtmt				0.004	16
Base. 2				0.052	16
<i>N</i> (groups)	16	15	16	16	

p values of Wilcoxon signed-rank tests between the treatment specified in the row and that in the column, for the given phase (baseline 1, treatment, or baseline 2). All tests use group-level data

significant if the sample were larger. However, when we allow some social context by allowing subjects to communicate, normative messages have an effect that is significantly larger than information or communication alone (though not significantly larger than normative messages alone). While the point estimate of the effect for the normative + communication treatment is more than the sum of the point estimates for the normative and communication treatments, we cannot reject the hypothesis that it is equal to that sum. For comparison, with information only, [Ferraro and Price \(2013\)](#) find a 1% decline in water usage and [Cutter and Neidell \(2009\)](#) find a 2.5–3.5% drop in freeway traffic; with normative messaging, [Allcott \(2011\)](#) and [Ayres et al. \(2013\)](#) find declines in power use of 2 and [Ferraro and Price \(2013\)](#) find a decline in water use of 4.8%.²¹ In our study, if we translate effects into percentages, information reduces appropriation by 0.549%, communication by 2.812%, normative messages by 5.233%, and normative messages with communication by 11.441%. Thus, our results for the information and normative treatments are at the low and high ends, respectively, of the ranges of results from relevant field experiments, and our result from the normative + communication treatment appears larger.

After the non-price interventions are removed, appropriation levels do rise again, but not all the way to the Nash prediction,²² and in the second baseline the non-price interventions

²¹ The CPRs in those studies are large enough to preclude the situation where resource user A tries to convince others to reduce their use to increase resource returns so that A can get larger gains from exploitation.

²² While the number of individual choices at 10 tokens is nontrivial, so one might fear that the Nash of 9 tokens is too close to that edge to allow discrimination, in [Table 1](#) we can see that there is sufficient statistical power to distinguish between differences much smaller than a token. Therefore, we are confident declaring

Table 3 Communication message choices

	Communication	Norm. + comm.
Stay at the same level in BLUE	39	37
Good job, guys, stay at the same level in BLUE! :)	45	43
Invest more in BLUE	6	3
Good job, guys, but invest more in BLUE! :)	4	6
YOU GUYS NEED TO INVEST MORE IN BLUE! :(3	4
Invest less in BLUE	56	68
Good job, guys, but invest less in BLUE! :)	61	30
YOU GUYS NEED TO STOP INVESTING SO MUCH IN BLUE! :(68	76
I don't know what we should do	16	17
I'm sorry	6	2
No message	32	50
Total	336	336

Cells show how many messages of each type were sent in each treatment

perform better than the subsidy. In particular, while post-treatment decay of cooperation is a problem for the subsidy treatment, it is less of a problem for the normative messaging treatment, and not a problem for normative messaging with communication; seen another way, the normative messaging treatment seems to have better persistence properties, particularly when communication is possible, while the subsidy seems to have much worse. However, the change in mean group appropriation from the first to the second baseline is not significantly different across treatments ($p > 0.2$ in all cases).

Our communication treatment has a much smaller effect than some work in the literature (e.g. [Ostrom et al. 1992](#)); however, that study, for example, allowed subjects 10 min of free-form face-to-face communication, which should be more effective than anonymously choosing messages from a pre-selected list. Indeed, limited communication has been quite unsuccessful at increasing cooperation in some situations ([Wilson and Sell 1997](#)). On the other hand, our communication treatment is the only case in which we do not nudge subjects to reduce their extraction, so it is notable that we do see a reduction even here.

In the treatments that allowed communication, [Table 3](#) shows that the most commonly chosen messages ask group members to decrease their appropriation of the CPR, and the next most commonly chosen messages encourage continued appropriation at the same level. There is some evidence that a negative tone is chosen more often when communication is combined with a normative message than if it is not (Wilcoxon rank-sum $p = 0.094$ for a individual-level test of average tone message, where neutral messages have a value of 0 and positive and negative tone messages have values of 1 and -1 respectively).

Footnote 22 continued

Baseline 2 behavior in the Subsidy treatment to be indistinguishable from the Nash. Further, since we observe group average contributions essentially moving from one Nash to another when the subsidy is removed, this behavior seems too much like profit-maximization to be an accident.

In a regression not shown (but available in online supplementary material), we find that the tendency to send a message that instructs group members to reduce their appropriation of the CPR is significantly greater if the group's appropriation in the previous period was higher, such that a one token increase in group appropriation makes a person 4–5% more likely to send such a message, and this is true whether or not normative messages are being sent. This implies that subjects understand the over-appropriation problem and want to use communication to encourage better behavior within their group. However, subjects whose own appropriation in the previous period was greater are no more, and may be less, likely to send a message telling others to reduce their appropriation—a one token increase in previous-round appropriation is associated with up to a 2% reduction in the likelihood of sending a message telling others to reduce appropriation. Therefore, it seems that the more rapacious subjects are not trying to disingenuously get others to reduce appropriation to increase their own returns. While this might at first glance seem like a missed strategic opportunity, according to other regressions not shown (but also available in supplementary material), the messages chosen do not significantly influence group appropriation levels.

Examining average appropriation levels ignores two key points. First, in the normative, information, and normative + communication treatments, different subjects within a treatment receive different actual treatments—some get a message about over-extraction, and others about optimal extraction, and still others about under-extraction. Second, as discussed above, an improvement in average group extraction need not imply increased efficiency since variation in appropriation across groups matters. We address these concerns in two ways. First, we break groups into over- and under-extractors and look at the effects each treatment has on the target groups separately. Second, we directly consider efficiency, which accounts for the nonlinear damages.

4.2 Extracting More or Less

Choices made by a group's members put that group in one of three categories: over-extractors (which appropriated more than the optimum), under-extractors (which appropriated less), and groups extracting at an optimal level. In the baseline phases and the subsidy treatment, all groups are treated identically, while in the information and normative treatments they are treated differently. The normative messaging and information treatments thus each contain three sub-treatments. Offsetting behavior by these different types could diminish the average effect of these treatments on appropriation while nonetheless increasing efficiency.

How these different group types behave has implications for efficiency. As discussed above, even if mean extraction is too high, when some groups are optimal extractors or under-extractors and the CPR production function is nonlinear, an across-group average reduction in extraction toward the social optimum can be anti-social on net. The efficiency of the policies, therefore, is not simply a matter of the average appropriation level but also the different effects the policies have on groups in these different categories.

In the information treatment, of the 112 (=7 rounds \times 16 groups) group messages transmitted in the treatment, 92 messages informed groups that a decrease in appropriation would improve group payoff, 14 that an increase would improve group payoff, and 6 that they were at the maximum group payoff. Consequently, 17.9% of the messages sent to subjects would not be expected to cause reduced appropriation. Similarly, of the 105 (=7 rounds \times 15 groups) group messages transmitted in the normative messaging treatment, 81 encouraged a decrease in appropriation, 13 encouraged an increase, and 11 informed subjects they were at an optimum. Here, 22.9% of messages would not be expected to reduce appropriation. With normative messaging and communication, out of 112 (=7 rounds \times 16 groups) messages, 82

encouraged a decrease, 21 an increase, and 9 no change in appropriation, so 26.8% should not reduce appropriation.

If we leave aside communication, an over-extracting group can see one of four treatments in a following period: nothing (baseline), a subsidy on the outside option (subsidy), information on the socially optimal action (information), or a normative appeal to decrease appropriation (normative). To estimate the effect of a normative message in reducing group over-extraction, we must compare later behavior of an over-extracting group that gets a normative message to the later behavior of this group when it receives no message and no other treatment. The treatment effect we would like to identify for such groups is:

$$E(\Sigma x_{jt} | \text{over-extractor}_{t-1}, \text{normative}) - E(\Sigma x_{jt} | \text{over-extractor}_{t-1})$$

If this difference in group-level appropriation is negative, then the normative message was effective. In the analogous construction for under-extractors, efficiency would increase with a positive treatment effect. We do not observe the same group in multiple treatments and conditions in a given round. However, treatment type is exogenously varied by the experimenter, and every group is observed both with the treatment and without (in the baseline rounds).²³ We can therefore estimate treatment effects using the model:

$$\Sigma x_{jt} = \beta_0 + \beta_1(\text{overex}_{t-1}) + \beta_2(\text{overex}_{t-1} \times \text{normative}) + \Gamma X_{jt} + \epsilon$$

Here, *overex*_{*t*-1} is an indicator for whether the group over-extracted in the last period, *X*_{*jt*} is a collection of group- and period-specific covariates, and Γ is the vector of coefficients for those covariates. The average treatment effect on the treated—the effect of the treatment on this type of group—is then β_2 . By constructing indicator variables for over-extractors, under-extractors, and socially-optimal extractors, as well as for each of the five treatments, and including full interactions with each treatment, we extend this model to isolate treatment effects for each treatment on each group type.²⁴

Table 4 presents regression estimates of treatment effects. Specification 1 is a simple pooled OLS regression. Specification 2 uses fixed effects and therefore can be interpreted as giving the average of within-group effects of each treatment on the total appropriation. In both Specifications 1 and 2, the second baseline periods are omitted, so that appropriation under treatment is being compared only to the first baseline rounds. Specification 3 and 4 use pooled OLS and fixed-effects models, respectively, and include all rounds, using dummies to control for the interaction between the second baseline and the treatment.^{25, 26} Specification 2 is our base specification for the discussion that follows.

First, we see that groups that over-extracted in the last period extract more than do groups that were optimal-extractors. Since this effect persists in within-group specifications, this may be caused by conditional cooperation rather than persistence of behavioral “types.”

²³ While treatment is exogenously varied, the valence of normative and information messages are endogenous to group behavior; therefore we cannot speak to the effects of “mismatched” messages, e.g. a message to under-extractors that they should reduce extraction.

²⁴ Without dummies indicating the group’s “type” (over-, under-, or optimal-extractor in the last period), the treatment-type interaction dummies would give effects relative to an inappropriate counterfactual that includes groups of other types, thus biasing results.

²⁵ These results are robust to other sets of control variables. Results are unchanged with random effects, or when a Tobit model (which accounts for group-level censoring in appropriation) is used.

²⁶ When random-effects regressions are conducted at the individual level, the number of times the subject used the tutorial has a significant and positive coefficient of 0.180–0.199. This seems to be driven by high appropriation levels by a small number of subjects who each practiced a large number of times. It is not possible to determine how much a subject uses the calculator available in the decision screen.

Table 4 Regressions of group-level appropriation on covariates

Specification	1 Pooled OLS	2 Fixed effects	3 Pooled OLS	4 Fixed effects
<i>Over-extractors</i>				
Over-extracting group in previous period	3.492*** (0.542)	2.890*** (0.528)	3.613*** (0.509)	2.893*** (0.460)
Interacted with subsidy	-7.075*** (0.692)	-7.908*** (0.518)	-7.043*** (0.683)	-7.727*** (0.545)
Information	-1.068* (0.552)	-0.785 (0.591)	-1.037* (0.552)	-0.744 (0.573)
Normative	-2.261*** (0.676)	-1.996*** (0.581)	-2.229*** (0.673)	-1.977*** (0.582)
Communication	-0.925* (0.520)	-1.306** (0.501)	-0.893* (0.514)	-1.359** (0.536)
Norm. + comm.	-3.592*** (0.635)	-3.187*** (0.809)	-3.561*** (0.634)	-3.200*** (0.815)
<i>Optimal extractors</i>				
Socially optimal group in previous period	Omitted	Omitted	Omitted	Omitted
Interacted with subsidy	-3.331*** (0.716)	-4.124*** (0.802)	-3.178*** (0.671)	-4.238*** (0.729)
Information	-0.478 (1.638)	1.608 (1.023)	-0.325 (1.630)	1.512 (0.998)
Normative	-0.534 (0.892)	0.781 (1.236)	-0.371 (0.877)	0.956 (1.170)
Communication	-0.423 (1.682)	0.87 (2.377)	-0.270 (1.457)	1.174 (1.840)
Norm. + comm.	-0.756 (1.306)	-0.089 (1.446)	-0.603 (1.296)	-0.023 (1.381)
<i>Under-extractors</i>				
Under-extracting group in previous period	-1.542* (0.831)	0.028 (0.712)	-0.899 (0.758)	0.083 (0.643)
Interacted with subsidy	-1.841** (0.831)	-3.780*** (0.956)	-2.330*** (0.724)	-3.965*** (0.909)
Information	-0.794 (1.751)	0.577 (1.317)	-1.283 (1.706)	0.342 (1.266)
Normative	0.41 (0.722)	0.086 (0.950)	-0.080 (1.116)	-0.172 (0.909)
Communication	0.064 (1.163)	-0.613 (1.295)	-0.426 (1.139)	0.061 (1.046)
Norm. + comm.	0.23 (0.993)	-1.037 (1.170)	-0.260 (0.970)	-1.049 (1.115)

Table 4 continued

Specification	1	2	3	4
	Pooled OLS	Fixed effects	Pooled OLS	Fixed effects
Second baseline	Omitted	Omitted	Dummies	Dummies
Constant	20.978*** (0.525)	21.264*** (0.486)	20.825*** (0.494)	21.256*** (0.448)
Observations ($N \times t$)	1106	1106	1659	1659
Overall R^2	0.312	0.279	0.368	0.342
Number of groups	79	79	79	79

Dependent variable is group appropriation in tokens (0–30). Spec. 1 and 2 omit second baseline rounds; Spec. 3 and 4 include dummies interacting each treatment dummy with the second baseline dummy. First baseline is the omitted category. Robust standard errors (clustered on group) in parentheses

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

Second, we discuss the effect of our benchmark price intervention, the subsidy. It had a very large effect on over-extraction. The subsidy had a less desirable outcome among the other types of groups. It led to subsequent under-extraction by optimally-extracting groups and it exacerbated existing under-extraction. It should be unsurprising that a greater incentive to select the outside option increases the use of that outside option regardless of the initial level of appropriation. Group-level heterogeneity in CPR use means that the subsidy can, in a significant minority of cases, have anti-social consequences.

All non-price policies also significantly reduced over-extraction relative to the baseline, although the information treatment’s effect is small and sensitive to specification. However, all non-price policies were significantly less effective at reducing over-extraction by over-extractors as compared to the subsidy ($p < 0.001$). Normative messaging with communication is also significantly more effective at reducing over-extraction than either communication alone ($p = 0.055$) or information ($p = 0.020$), but there are no other significant differences in coefficients for over-extractors among the nonpecuniary treatments.

The effects of the nonpecuniary treatments were significantly different from the effect of the subsidy when applied to optimally- or under-extracting groups ($p < 0.100$ in nearly all cases²⁷). In very few cases were the nonpecuniary treatments significantly different from one another: in the pooled OLS models, the effect of normative messaging with communication differed from that of the information treatment ($p = 0.003$ in both specifications) as well as that of the communication treatment ($p < 0.001$ in both specifications). The nonpecuniary treatments did not have a statistically significant effect on appropriation by under- or optimally-extracting groups, which means that while they did not cure under-extraction, unlike the subsidy, they did no harm.

In results not shown here but available in the online supplementary material, we find in Specifications 3 and 4 that the second baseline sees significantly more extraction than the first baseline in the subsidy treatment ($p < 0.001$) and perhaps in the information treatment ($p = 0.112$ and $p < 0.001$ in Specifications 3 and 4 respectively) and the communication treatment ($p = 0.027$ and $p = 0.101$, respectively) but not in the normative ($p = 0.851$ and $p = 0.419$, respectively) or normative with communication ($p = 0.869$ and $p = 0.513$, respectively) treatments. This echoes our previous finding that normative messaging,

²⁷ In Specification 1, the effect of information was not significantly different from the effect of the subsidy on under-extractors, $p = 0.514$, and was almost significantly different in Specification 3, $p = 0.101$.

Table 5 Efficiency levels by treatment

	Pooled	Subsidy	Information	Normative	Communication	Norm. + comm.
Base. 1	0.859 (0.074)	0.839 (0.070)	0.883 (0.076)	0.863 (0.082)	0.842 (0.064)	0.869 (0.076)
Treatment	0.920 (0.068)	0.980 (0.014)	0.884 (0.074)	0.914 (0.072)	0.885 (0.069)	0.938 (0.014)
Base. 2	0.819 (0.081)	0.753 (0.050)	0.822 (0.084)	0.856 (0.068)	0.806 (0.087)	0.861 (0.068)
<i>N</i> (groups)	47	16	16	15	16	16
Within-treatment test <i>p</i> values						
Base. 1 = Treatment	<0.001	0.001	1.000	0.023	0.017	0.006
Treatment = Base. 2	<0.001	<0.001	0.001	0.003	0.001	0.001
Base. 1 = Base. 2	<0.001	0.002	0.001	0.570	0.044	0.756

Standard deviations in parentheses. Values in percent of maximum possible CPR yield achieved (0–1). Within-treatment *p* values are of Wilcoxon signed-rank tests. All tests use group-level data

particularly with communication, has the best persistence properties, that is, that it shows the least decay in cooperation after treatment is removed.

To summarize, we have reproduced the effectiveness that Pigouvian instruments have shown in existing literature. However, we are the first to show that the subsidy causes not only over-extractors but also under-extractors to reduce their activity, and this latter effect is an exacerbation of existing socially suboptimal behavior. We have also provided the first direct comparison of the effectiveness of several commonly-discussed nonpecuniary policies to the benchmark of no intervention and that of a Pigouvian intervention. All have some ability to reduce over-extraction, although that effect is much weaker than the Pigouvian instrument's effect, and the strongest of these is a normative message when communication is allowed within the group. We also show that the gains of the nonpecuniary treatments are not offset by efficiency losses resulting from reduced extraction by under-extracting groups.

4.3 Achieving and Maintaining Economic Efficiency

The subsidy is more effective than the nonpecuniary treatments in fighting over-extraction; however, the subsidy's increased efficiency is offset by efficiency losses from reduced extraction by optimal extractors and under-extractors, which does not occur with the other treatments. On net, how does the efficiency of the non-price treatments compare to that of the subsidy?

Earlier we defined efficiency as the actual return from the CPR divided by the return from the CPR at the social optimum. Table 5 presents comparisons of efficiency across treatments and against the baseline periods, and Table 6 shows additional test statistics.²⁸

Again, the first baselines do not differ across the treatments, as they should not.

The subsidy yielded the highest efficiency during treatment, reaching 98.0% of potential resource yield.²⁹ However, after the subsidy is removed, appropriation reaches the least efficient level in any of our data, at 75.3%.

²⁸ The results are robust to a panel regression specification.

²⁹ The efficiency of the subsidy is higher than the 93% observed with a Pigouvian tax in Plott (1983), but Plott notes that the efficiency of his tax treatment was likely not statistically different from that of his license

Table 6 Tests of efficiency levels across treatments

	Information	Normative	Communication	Norm. + comm.	N (groups)
Subsidy					
Base. 1	0.127	0.363	0.955	0.175	16
Trtmt	<0.001	<0.001	<0.001	<0.001	16
Base. 2	0.008	<0.001	0.050	<0.001	16
Info.					
Base. 1		0.693	0.163	0.429	16
Trtmt		0.260	0.985	0.036	16
Base. 2		0.236	0.429	0.214	16
Norm.					
Base. 1			0.286	0.906	15
Trtmt			0.173	0.441	15
Base. 2			0.058	0.843	15
Comm.					
Base. 1				0.228	16
Trtmt				0.016	16
Base. 2				0.035	16
N (groups)	16	15	16	16	

Standard deviations in parentheses. Between-treatment *p* values are of Wilcoxon rank-sum tests; within-treatment *p* values are of Wilcoxon signed-rank tests. All tests use group-level data

By comparison, all of the non-price treatments also increase efficiency in some sense: all except the information treatment see an increase in efficiency from the first baseline to the treatment period, and all see an efficiency drop when they are removed. Of the non-price interventions, normative messaging with communication performed best, at 93.8 % efficiency. During treatment, all non-price interventions were significantly less efficient as compared to the subsidy. Normative messaging with communication is more efficient than either the information or communication treatments, while all other comparisons are not statistically significant. However, because deadweight loss increases with the square of the distance from the social optimum, the efficiency gains from even the small reductions in over-extraction in the nonpecuniary treatments were large.

When treatment ceases, efficiency drops significantly in all cases. However, as compared to the drop observed when the subsidy is removed, this decay is smaller after normative messaging (with or without communication) stops. In fact, the normative treatments are the only cases in which efficiency did not decline between the first and second baselines. The decay from first to second baseline is less in the normative treatment than in either the subsidy (Wilcoxon rank-sum $p = 0.013$) or the information ($p = 0.069$) treatments, and also less in the normative treatment with communication than in the subsidy treatment ($p = 0.019$), while there are no other significant across-treatment differences in baseline efficiency shift.

Footnote 29 continued treatment (at 98.3 %). If there is a difference, it may be due to tax aversion (see, e.g. Kallbekken et al. 2011 and Heres et al. 2013).

Either the normative messages forestalled a decay in efficiency that would have happened without any intervention, or the subsidy had harmful effects on later cooperation. The latter interpretation is consistent with findings that extrinsic motivation frames social dilemmas as transactional, undermining intrinsic social motivation and causing motivational crowding out (e.g. [Gneezy and Rustichini 2000](#); [Frey and Jegen 2001](#)).

Are these differences in efficiency more attributable to differences in under-extraction or over-extraction? A closer look at the data reveals that inefficiency caused by over-extractors is always at least 25 times that caused by under-extractors, sometimes much more, except during treatment by the subsidy. In that case, under-extractors cause more than twice as much inefficiency (1.4 % efficiency loss) as compared to over-extractors (0.6 % loss).

We have shown that CPR extraction under the subsidy treatment is more efficient than the nonpecuniary treatments. The efficiency analysis presented above, however, accounts only for efficiency of resource use, without regard for additional sources of efficiency loss. The subsidy faces the largest such problems, the most significant of which is that the government must raise revenues to pay a subsidy. If efficiency costs of raising public funds are considered, the subsidy becomes less efficient. Therefore, which policy is truly most efficient depends on the marginal cost of public funds. At reasonable estimates of the marginal cost of public funds, the policies considered here are indistinguishable on efficiency grounds.³⁰ Policies also require administration. It seems likely that the administrative costs of the information and normative policies are at worst no greater than the administrative costs of a subsidy because Pigouvian instruments would require monitoring and enforcement at the level of individual users while the other interventions would not.

5 Conclusion

CPR problems large and small bedevil environmental management. Environmental economics has offered Pigouvian taxes and subsidies as first-best tools to correct externalities, but they are naturally difficult to implement in CPR settings where monitoring and enforcement may be impossible. We use a lab testbed to compare, on multiple dimensions, a Pigouvian subsidy to nonpecuniary interventions that have gained policy attention in recent years: information provision, communication, and information with a normative appeal. Our contribution to existing literature is that our lab environment allows us to make careful comparisons among several non-price interventions and between these interventions and a theoretically-efficient price intervention, and allows us to evaluate performance on several metrics: average extraction, performance with both over- and under-extractors, efficiency, and post-intervention persistence.

We show that the price power of the subsidy has a large effect as predicted by theory, but has some unappealing properties: it causes groups that are inefficiently under-using a resource to further under-use (increasing inefficiency), and it has poor persistence when the subsidy is removed. Nonpecuniary interventions cause smaller but significant decreases in appropriation: information reduces appropriation by 0.549 %, communication by 2.812 %, normative messages by 5.233 %, and normative messages with communication by 11.441 %. Most of these values are of the same order of magnitude as the effects observed in field experiments that study the effects of information and normative messages on actual conservation behavior (e.g. [Ferraro and Price 2013](#); [Ayres et al. 2013](#)), although our normative message treatment

³⁰ For example, [Ballard et al. \(1985\)](#) find that each additional dollar of revenue comes at a welfare cost of 17–56 cents, and [Allgood and Snow \(1998\)](#) estimate the cost to be 13–28 cents.

with communication outperforms those results. This is so even though a social context, in which peer pressure and reputation may act in concert with policy, is deliberately excluded from our study, and the limited communication allowed in our communication treatments only provide a small amount of social context.

All nonpecuniary tools only change the behavior of groups that had previously been over-using their resource. The addition of “best practices”-type normative language has no significant effect beyond the effect of simple informational messages, and allowing communication does not significantly increase the effect of normative messages. Of all policies we tested, however, normative messages with and without communication have the best persistence properties. While the subsidy is more efficiency-enhancing than the nonpecuniary policies in this context, a subsidy in a non-lab situation would incur efficiency costs through the cost of raising public funds and through the added monitoring and enforcement needed to implement a system that targets individual users (as the subsidy must) rather than groups (as nonpecuniary treatments do). Our results, therefore, put the size of results from existing field experiments into context, showing that while the effects of non-price interventions are small, efficiency gains may be large relative to meaningful counterfactuals.

These results are important because around the world, governments and non-governmental entities are trying to identify the best policies for addressing the CPR problems that face their constituencies. The context of our lab experiment differs from the contexts of actual CPRs of concern, so one cannot translate our results directly to policy. However, our results can shed light on some results from field experiments and observational data, and can give some behavioral insights about comparative properties of these policies in the abstract. Price instruments have been tried, from Ireland’s green taxes to the United States’s hybrid car tax credit, and these are inarguably effective. However, these instruments are not always possible. They would be physically infeasible for some resources, such as open access fisheries, while in other cases they are simply politically infeasible—and it is fair to expect that price instruments at truly Pigouvian levels (levels that would actually correct the externality) will be rare in the world for the foreseeable future. Nonpecuniary strategies appear less powerful on first blush, but we have shown that when considered holistically, their efficiency properties compare well to a subsidy.

Unlike subsidies, taxes do not suffer from the efficiency costs of public revenue-raising and in fact may provide a double dividend (Goulder 1995), and thus they would compare more favorably to the nonpecuniary instruments we study. However, perhaps because of tax aversion (e.g. Kallbekken et al. 2011), taxes can face intense resistance, while policymakers have shown a taste for green subsidies in many settings (e.g. the United States Conservation Reserve Program and Renewable Electricity Production Credit).

It is likely that CPR problems in different contexts will be best served by different interventions. For example, climate change presents challenges of such great scale that a price intervention might be the best global solution to reduce greenhouse gas emissions. More research is needed on the interaction between policy effectiveness and scale. However, our results remind us that even in such settings, we need to consider the human element: the Pigouvian price calculated assuming self-interested behavior is not the right price for everyone, and adjustments should be made for heterogeneous behavior across groups. This heterogeneity is particularly important when resources are small in scale, as in the case of local parks. If people sort according to preferences, heterogeneity across the groups that share distinct small resources will intensify, and the need for policy to treat under-users differently from over-users becomes greater in such cases. Also in small-scale resources, since one person’s restraint increases returns for others, some users may strategically take advantage of others’

moderation by increasing their extraction. This is a phenomenon we do not find evidence of in our current study but that merits future research.

Innovative policy solutions may seek the best of both worlds by combining pecuniary instruments with nonpecuniary interventions. For example, Ireland's tax on plastic bags was accompanied by a major publicity campaign and substantial efforts to get buy-in from stakeholders, and it is believed that this is in part why it achieved such impressive results (Convery et al. 2007). As we have shown and as shown by Allcott and Rogers (2014) in the context of electricity use and Ferraro et al. (2011) in the context of water conservation, nonpecuniary policies can have good persistence properties, and our results show that these persistence properties are much better than those of the price instrument. It is possible that a combination policy will have the efficiency properties of the price instrument without the crowding-out that might normally accompany such a policy. Future research is needed to study complementarities that may arise when multi-pronged policies are implemented to manage CPR problems.

References

- Allcott H (2011) Social norms and energy conservation. *J Public Econ* 95(910):1082–1095 Special Issue: The Role of Firms in Tax Systems
- Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation. *Am Econ Rev* 104(10):3003–3037
- Allgood S, Snow A (1998) The marginal cost of raising tax revenue and redistributing income. *J Polit Econ* 106(6):1246–1273
- Anderson, L, Holt CR, Reiley D (2007) Congestion pricing and welfare: an entry experiment. In: Cherry TL, Kroll S, Shogren JF (eds) *Environmental economics, experimental methods*. Routledge, pp 280–293
- Andreoni J (1989) Giving with impure altruism: applications to charity and Ricardian equivalence. *J Polit Econ* 97:1447–1458
- Arrow KJ (1984) The economics of agency. In: Richard Z, John WP (eds) *Principals and agents: the structure of business*. Harvard Business School Press, Boston
- Ayres I, Raseman S, Shih A (2013) Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *J Law Econ Organ* 29(5):992–1022
- Ballard CL, Shoven JB, Whalley J (1985) General equilibrium computations of the marginal welfare costs of taxes in the United States. *Am Econ Rev* 75(1):128–138
- Brown G (1974) An optimal program for managing common property resources with congestion externalities. *J Polit Econ* 82(1):163–173
- Bru L, Cabrera S, Capra CM, Gomez R (2003) A common pool resource game with sequential decisions and experimental evidence. *Exp Econ* 6(1):91–114
- Budescu DV, Suleiman R, Rapoport A (1995) Positional order and group size effects in resource dilemmas with uncertain resources. *Organ Behav Hum Dec Process* 61(3):225–238
- Casari M, Plott CR (2003) Decentralized management of common property resources: experiments with a centuries-old institution. *J Econ Behav Organ* 51(2):217–247
- Cochard F, Willinger M, Xepapadeas A (2005) Efficiency of nonpoint source pollution instruments: an experimental study. *Environ Resour Econ* 30(4):393–422
- Coleman EA (2009) A comparison of demand-side water management strategies using disaggregate data. *Public Works Manag Policy* 13(3):215–223
- Convery F, McDonnell S, Ferreira S (2007) The most popular tax in Europe? Lessons from the Irish plastic bags levy. *Environ Resour Econ* 38(1):1–11
- Cox JC, Deck CA (2005) On the nature of reciprocal motives. *Econ Inq* 43(3):623–635
- Cox JC, Friedman D, Gjerstad S (2007) A tractable model of reciprocity and fairness. *Games Econ Behav* 59(1):17–45
- Cox JC, Ostrom E, Walker JM, Castillo A, Coleman E, Holahan R, Schoon M, Steed B (2009) Trust in private and common property experiments. *South Econ J* 75(4):957–975
- Cutter WB, Neidell M (2009) Voluntary information programs and environmental regulation: evidence from spare the air. *J Environ Econ Manag* 58(3):253–265
- Dawes RM (1980) Social dilemmas. *Annu Rev Psychol* 31(1):169–193

- Dulleck U, Kaufmann S (2004) Do customer information programs reduce household electricity demand? The Irish program. *Energy Policy* 32(8):1025–1032
- Falk A, Fehr E, Fischbacher U (2002) Appropriating the commons: a theoretical explanation. In: Ostrom E, Dietz T, Dolsak N, Stern PC, Stonich S, Weber EU (eds) *The drama of the commons*. National Academy Press, Washington, pp 157–191
- Ferraro PJ, Price MK (2013) Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. *Rev Econ Stat* 95(1):64–73
- Ferraro PJ, Jose Miranda J, Price MK (2011) The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment. *Am Econ Rev* 101(3):318–322
- Festinger L (1954) A theory of social comparison processes. *Hum Relat* 7:117–140
- Frey BS, Jegen R (2001) Motivation crowding theory. *J Econ Surv* 15(5):589–611
- Gächter S (2007) Conditional cooperation: behavioral regularities from the lab and the field and their policy implications. In: Frey BS, Stutzer A (eds) *Psychology and economics: a promising new cross-disciplinary field (CESifo seminar series)*. The Mit Press, Cambridge, pp 19–50
- Gneezy U, Rustichini A (2000) A fine is a price. *J Leg Stud* 29:1
- Goeree JK, Holt CA, Laury SK (2002) Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *J Public Econ* 83(2):255–276
- Goldstein NJ, Cialdini RB, Griskevicius V (2008) A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *J Consum Res* 35(3):472–482
- Goulder LH (1995) Environmental taxation and the double dividend: a reader's guide. *Int Tax Public Finance* 2(2):157–183
- Harrison GW, Hoffman E, Rutstrom EE, Spitzer ML (1987) Coasian solutions to the externality problem in experimental markets. *Econ J* 97(386):388
- Haynie AC, Hicks RL, Schnier KE (2009) Common property, information, and cooperation: commercial fishing in the bering sea. *Ecol Econ* 69(2):406–413
- Heres D, Kallbekken S, Galarraga I (2013) Understanding public support for externality-correcting taxes and subsidies: a lab experiment. Working paper 2013–04, BC3
- Janssen MA, Anderies JM, Joshi SR (2011) Coordination and cooperation in asymmetric commons dilemmas. *Exp Econ* 14(4):547–566
- Johansson O (1997) Optimal Pigovian taxes under altruism. *Land Econ* 73(3):297
- Kallbekken S, Kroll S, Cherry TL (2011) Do you not like Pigou, or do you not understand him? Tax aversion and revenue recycling in the lab. *J Environ Econ Manag* 62(1):53–64
- LaRiviere J, Holladay S, Price MK, Novgorodsky D (2014) Prices vs. nudges: a large field experiment on energy efficiency fixed cost investments. Technical report, working paper
- Marwell G, Ames R (1979) Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem. *Am J Sociol* 84(6):1335–1360
- Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393(6685):573–577
- Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437(7063):1291–1298
- Osés-Eraso N, Viladrich-Grau M (2011) The sustainability of the commons: giving and receiving. *Exp Econ* 14(4):458–481
- Ostrom E (1992) The rudiments of a theory of the origins, survival, and performance of common property institutions. Theory, making the commons work: practice and policy
- Ostrom E (1998) A behavioral approach to the rational choice theory of collective action: presidential address, American Political Science Association, 1997. *Am Polit Sci Rev* 92(1):1–22
- Ostrom E (2000) Reformulating the commons. *Swiss Polit Sci Rev* 6(1):29–52
- Ostrom E (2002) Chapter 24 common-pool resources and institutions: toward a revised theory. In: Bruce LG, Gordon CR (eds) *Handbook of agricultural economics*, vol 2, part A. Elsevier, pp 1315–1339
- Ostrom E, Walker J, Gardner R (1992) Covenants with and without a sword: self-governance is possible. *Am Polit Sci Rev* 86(02):404–417
- Ostrom E, Gardner R, Walker J (1994) *Rules, games, and common-pool resources*. University of Michigan Press, Ann Arbor
- Pigou AC (1920) *The economics of welfare*. McMillan, London
- Plott CR (1983) Externalities and corrective policies in experimental markets. *Econ J* 93(369):106
- Renwick ME, Green RD (2000) Do residential water demand side management policies measure up? An analysis of eight California water agencies. *J Environ Econ Manag* 40(1):37–55
- Rodriguez-Sickert C, Andrés Guzmán R, Camilo Cárdenas J (2008) Institutions influence preferences: evidence from a common pool resource experiment. *J Econ Behav Organ* 67(1):215–227
- Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V (2007) The constructive, destructive, and reconstructive power of social norms. *Psychol Sci* 18(5):429–434

- Suter JF, Duke JM, Messer KD, Michael HA (2012) Behavior in a spatially explicit groundwater resource: evidence from the lab. *Am J Agric Econ* 94(5):1094–1112
- Velez MA, Stranlund JK, Murphy JJ (2009) What motivates common pool resource users? Experimental evidence from the field. *J Econ Behav Organ* 70(3):485–497
- Vossler CA, Poe GL, Schulze WD, Segerson K (2006) Communication and incentive mechanisms based on group performance: an experimental study of nonpoint pollution control. *Econ Inq* 44(4):599–613
- Walker JM, Gardner R, Ostrom E (1990) Rent dissipation in a limited-access common-pool resource: experimental evidence. *J Environ Econ Manag* 19(3):203–211
- Wilson RK, Sell J (1997) “Liar, Liar..”: cheap talk and reputation in repeated public goods settings. *J Confl Resol* 41(5):695–717