Journal of Environmental Economics and Management 98 (2019) 102276

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

Graduated stringency within collective incentives for group environmental compliance: Building coordination in field-lab experiments with artisanal gold miners in Colombia

Luz A. Rodriguez [a, *], Alexander Pfaff [b], Maria Alejandra Velez [c]

[a] School of Environmental and Rural Studies, Pontificia Universidad Javeriana, Colombia
[b] Sanford School of Public Policy, Duke University, United States
[c] School of Economics and Center for the Study of Security and Drugs (CESED), Universidad de Los Andes, Colombia

A R T I C L E I N F O

Article history:
Received 16 July 2018
Received in revised form 5 October 2019
Accepted 30 October 2019
Available online 5 November 2019

JEL classification:
D70
C92
Q58

Keywords:
Collective incentives
Environmental policy
Graduated stringency
Field-lab experiments
Small-scale gold mining
Colombia

A B S T R A C T

Small-scale gold mining is important to rural livelihoods in the developing world but also a source of environmental externalities. Incentives for individual producers are the classic policy response for a socially efficient balance between livelihoods and the environment. Yet monitoring individual miners is ineffective, or it is very costly, especially on frontiers with scattered small-scale miners. We ask whether monitoring at a group level effectively incentivizes cleaner artisanal mining by combining lower-cost external monitoring with local collective action. We employ a mining-framed, threshold-public-goods experiment in Colombia's Pacific region, with 640 participants from frontier mining communities. To study compliance with collective environmental targets, we vary the target stringency, including to compare increases over time in the stringency versus decreases. We find that collective incentives can induce efficient equilibria, with group compliance – and even inefficient overcompliance – despite the existence of equilibria with zero contributions. Yet, for demanding targets in which the reward for compliance barely outweighs the cost, compliance can collapse. Those outcomes improve with past successes for easier targets, however, so our results suggest gain from building coordination via graduated stringency.

1. Introduction

Small-scale gold mining (SSGM) is an important economic activity within the developing world, including for many poor rural households living far out on the frontier. Yet it has many negative environmental consequences (Hinton, 2005; Hentschel et al., 2003; Hilson, 2003). It is the world's largest mercury polluter, putting ~100 million people at risk (Wade, 2013), and a leading driver of landscape change in environmentally critical regions such as the Amazon (Asner et al., 2013; Swenson et al., 2011). Alluvial mining removes significant quantities of sand, gravel and rock, contributing to sedimentation and acidification of important rivers that, in turn, carry emissions downstream, sometimes to large populations. In addition, open pits filled with water can be sources of mosquito-borne diseases like malaria (Castellanos et al., 2016; Crompton et al., 2002).

Regulation has not addressed these issues on the frontier. Command-and-control policies suitable for interactions with large firms have been applied to small-scale mining, despite states' inability to enforce for scattered small-scale miners

* Corresponding author. Pontificia Universidad Javeriana, Transversal 4 No 42-00 Piso 8, Edificio José Rafael Arboleda, Bogotá, Colombia.
E-mail address: rodriguez.luza@javeriana.edu.co (L.A. Rodriguez).

https://doi.org/10.1016/j.jeem.2019.102276
0095-0696/

(Hinton, 2005). This has yielded chaotic, open-access, poorly organized mining (Shoko, 2002). Further, unlike for common-pool resources where only local users are affected negatively by neighbors' resource appropriations (e.g., Ostrom, 1990), SSGM generates distant externalities that do not motivate local communities to act collectively to reduce environmental impacts (Saldarriaga-Isaza et al., 2013). Local rights alone, then, will not necessarily internalize most externalities from mining. Yet, local interaction may aid governance.

Given the strengths and the weaknesses of state and local institutions for addressing such issues, we used a decision experiment in the field to explore the impacts of a collective-rewards policy, for which success depends on local interactions. Incentives are external, due to states or NGOs, yet rewards are conditioned on group compliance,[1] using group metrics more feasible or cheaper for non-locals to monitor. Thus, incentives are focused not on individuals but instead on groups (see Kotchen and Segerson, 2019). Collective incentives create a local public good for such groups, so successes depend on the group's ability to solve a policy-induced collective-action problem.[2] Thus, pre-existing local norms, social capital, and prior interactions should affect the outcomes. Even if miners are hard to monitor, observable aggregates (e.g., forest cover or water quality) may still provide a sufficient basis for incentives which spur group interactions and compliance.

Our experimental design – drawing upon frontier fieldwork with miner interviews and surveys – is a threshold-public-good experiment with 640 inhabitants of frontier mining communities in Colombia's Pacific region, one of the country's poorest and also most biodiverse regions. We study group compliance and individual behaviors, while varying the stringency of aggregate or group environmental targets. We compare different stringency sequences, specifically sequences with falling stringency versus rising or graduated stringency. We find that collective rewards can induce efficient group compliance, and even inefficient overcompliance, despite the existence of an equilibrium involving no contributions. For demanding targets, in which collective rewards barely outweigh compliance costs, coordination and compliance can collapse completely. Yet, if a group has succeeded previously, facing an easier target, that improves the outcomes for more demanding targets. This implies the importance of building coordination via prior successful local interactions, supporting graduated stringency in the design of such incentives. Our results indicate that such collective environmental incentives can support environmental-and-economic local 'win-wins', based on lower-cost aggregate monitoring that can spur local collective action.

The paper is organized as follows. Section 2 discusses relevant literature, both theoretical and experimental. Section 3 describes our local setting, and Section 4 our experimental design and implementation. Section 5 presents our results and, finally, Section 6 provides further discussion.

2. Relevant literature

2.1. Theory concerning collective incentives with environmental targets

When it is difficult or impossible to observe individual behavior or to infer it from the outcomes, and outcomes depend on joint decisions by a group, there exists a moral hazard (Holmstrom, 1982). Such incentives to free ride arise in numerous situations such as teamwork, contributions to public goods (e.g., paying taxes), the management of common-pool-resources and non-point-source pollution. Each features a difference between individually and socially optimal actions. Some institution, exogenous or endogenous, is required to induce the socially optimal decisions.

Proposed institutions include incentives based on group performance. Levi (1988), for example, argues that states can induce citizens to pay taxes through coercion ("compulsory compliance"), social norms ("voluntary compliance"), or ideology backed by the use of coercion should a target not be collectively met ("quasi-voluntary", conditional compliance). Even people willing to pay voluntarily may require assurance that others also will pay in order for them to make payments. Public goods based on group compliance can induce 'quasi-voluntary' coordinated compliance.

Despite relatively little attention to group performance in the environmental economics literature, groups feature in water quality, air quality, fisheries, and land use (Kotchen and Segerson, 2019). Non-point pollution in particular has been a key focus for the theoretical development of group-based policies.[3] A seminal paper by Segerson (1988) proposes a policy for ambient quality with a fixed fine, or subsidy, plus an incentive proportional to the difference between pollution and, for example, optimal emissions.[4] This is analogous to Holmstrom (1982), where free riding on team or joint production can be removed using incentives for collective compliance. In these schemes, a principal administers incentives that break the

[1] Conditioning outcomes upon measurable aggregate behaviors follows other collective liability institutions such as microcredit in developing countries – where banks cannot monitor households (Armendáriz de Aghion and Morduch, 2005) – and related designs for moral hazard in groups (Holmstrom, 1982), for non-point source pollution (Segerson, 1988; Xepapadeas, 1991) and for threshold public goods (e.g., Bagnoli and McKee, 1991; Cadsby and Maynes, 1999).

[2] Using rewards, as for poor developing frontiers, sanctions might yield social unrest. SSGM is a subsistence activity for most rural poor, thus it is hard to expect them to incur additional cost based on their choices. Collective rewards can include financial transfers, as within PES, official permits for miners in a village, certifications for associations (Fairmined Standard) or funding for local development projects such as for roads, schools, and productive activities.

[3] Kotchen and Segerson (2019) discusses several types of group-performance policies in a unifying framework. Examples include entire industries threatened with costly regulations or taxes if they fail to meet pollution control targets voluntarily; regulations to protect air or water based on ambient rather than individual pollution; payments for ecosystem services made to groups or communities; total allowable catch of bycatch limits, among others.

[4] Different values for parameters within this policy generate the different options for collective incentives instruments: a pure ambient tax-and-subsidy scheme, a pure ambient tax scheme, a group fine (or subsidy), or a mixed scheme.

budget-balance constraint.[5] Xepapadeas (1991) adds to this toolbox a random fine with subsidies, which can avoid breaking the budget balance.

A key feature of group-performance policies is that they create a local public good for the group, since getting the reward or avoiding the sanction depends on group interactions and expectations. If expecting others to contribute, one may comply. This may be pivotal in collective compliance. Yet, if others contribute enough, then one might prefer to free ride. Group responsibility for such collective targets may induce agents to jointly ensure aggregate standards are met voluntarily (Segerson, 1998), with incentives for peer monitoring and sanctioning (Miceli and Segerson, 2007). This suggests complementarities between "top-down" group policies and "bottom-up" incentives for self-governance, similar to common-pool-resource management (Kotchen and Segerson, 2019).

Group policies can be voluntary, mandatory, or a combination of both (Segerson, 1999; Segerson and Wu, 2006; Suter et al., 2010). In mandatory settings, the compliance burden falls on the agents. Regulators can impose fines or taxes for failure to meet a target. In voluntary settings, the burden is on regulators to induce participation (e.g., PES or green certifications) (Segerson, 2013, 1999). Voluntary and mandatory approaches can be complements, as not meeting environmental goals voluntarily can trigger taxes or sanctions (Segerson, 1999; Segerson and Wu, 2006). Such taxes can be fixed (Segerson and Wu, 2006) or vary with the degree of non-compliance (Suter et al., 2010). Appropriate targets and incentives can induce cost-minimizing abatement without imposing a tax due to the threat of the tax (Kotchen and Segerson, 2019; Segerson and Wu, 2006). Yet, this need not be the only subgame-perfect Nash equilibrium. Compliance can also be reached at higher costs. Given heterogeneous firms, some obtain a net gain of voluntary participation, which can induce self-enforcing equilibria[6] that do involve some level of free riding (Dawson and Segerson, 2008).

In sum, success for group-based policies depends on how the policies are designed, i.e., external incentives, and how well groups function, i.e., an ability to act collectively (Kotchen and Segerson, 2019).[7] Factors that affect the performance of collective incentives have been studied empirically mainly through the use of lab experiments. Below, we summarize main topics and conclusions.

2.2. Experiments for coordination games (threshold public goods and ambient instruments)

Collective incentives transform a social dilemma into a form of coordination problem. However, while they can make social gains achievable within subgame-perfect Nash equilibria, lump-sum group fines or rewards can also be subject to coordination failure due to multiple Nash equilibria. The achievement of positive solutions to such coordination problems depends on expectations. For instance, using a lab experiment, Barrett and Dannenberg (2017) find a significant proportion of groups end up facing a prisoner's dilemma situation despite having the opportunity to switch to playing a tipping (coordination) game. In light of the game's design, and how expectations were formed, it seems those participants who choose to play the coordination game had lower expectations about the potential for cooperation within a prisoner's dilemma game and also had higher expectations about coordination on mutually preferred equilibria within a 'tipping' game.

Factors that affect expectations and thus outcomes in coordination games include: dimensions of policy (stringency, incentive type), information, learning, and uncertainty. Lab experiments used to study collective incentives can be divided into those exploring ambient instruments (see, for instance, Cochard et al., 2005; Alpízar et al., 2004; Spraggon, 2002) and those studying contribution to threshold public goods (e.g., Barrett & Dannenberg 207; Dannenberg et al., 2014; Bagnoli and McKee, 1991; Suleiman and Rapoport, 1992; Cadsby and Maynes, 1999). Among the results relevant to our study are that higher thresholds raise contributions yet also the chance that a target will not be reached (Ledyard, 1995; Suleiman and Rapoport, 1992; Cadsby and Maynes, 1999). Uncertainty and ambiguity regarding thresholds hinder cooperation, making initial contributions key to group performance (Dannenberg et al., 2014; Barrett and Dannenberg, 2012). Learning has an ambiguous effect on efficiency. Some studies find deviations from efficiency fall with experience (Alpízar et al., 2004). Others find the opposite (Suleiman and Rapoport, 1992). While communication can help to achieve such coordination, it can also lead to strategic overcompliance due to collusion when facing marginal tax or subsidy policies (Vossler et al., 2006; Poe et al., 2004; Suter et al., 2008).

Most such experiments have been conducted with students. Some had field subjects. Alpízar et al. (2004) test for differences between mill managers and students and they reject equal behavior. Reichhuber et al. (2009) test Segerson, 1998's non-point mechanisms in a common-pool resource experiment with Ethiopian farmers whose harvest of non-timber forest

[5] The budget-balance constraint refers to the limit imposed by the sum of society members' valuations, or resources. Within the group-performance scheme, an external actor (the principal) could charge fines or taxes, or could provide subsidies or rewards, that shift what is available to society such that total payments differ from society's valuation.

[6] In this self-enforcing equilibrium: (1) each participant firm earns a profit at least as high as it would have earned if no firms participate; and (2) no firm will be able to benefit by unilaterally changing its decision to join or to leave.

[7] Group-based approaches have also been proposed, as well as implemented, for small loans to poor households who lack collateral (Armendáriz de Aghion and Morduch, 2005). Borrowers can be guarantors for each other: if a group member defaults, all members are denied subsequent loans. The roots of such microfinance are in ROSCAs (rotating savings and credit associations), informal financial institutions found mainly in developing countries and immigrant communities in the developed world (Besley et al., 1993). Social ties, trust, peer monitoring are central to the functioning of these schemes and help to explain re-payment rates (see, for instance, Stiglitz, 1990; Karlan, 2007; Hadi and Kamaluddin, 2015, Griffin and Husted, 2015; Feigenberg et al., 2013; Feigenberg et al., 2014).

products is not observed. They conclude that to a great extent initial decisions determine behaviors in all following periods (regardless of the instrument used) and that a high-tax mechanism is more effective in achieving the target than a tax-subsidy mechanism, mainly because the latter could generate tacit collusion. Saldarriaga-Isaza et al. (2015), in the context of SSGM and a threshold-public-goods experiment, examine 'associative entrepreneurship', or the creation of associations of small-scale miners, as a way to access costly environmentally friendly technologies. They find that miners' contributions alone were not sufficient to acquire the technologies. However, when there existed interventions by third-party advisors, then participants were able to achieve efficient levels of contributions.

Given the concern that students' behaviors cannot be counted on to equal those in the field, our experiments are conducted with members of small-scale gold-mining communities in Colombia. Our main contribution is to vary the stringency of targets, in particular the order of stringency: we compare policies that increase the stringency of targets with policies that decrease stringency.

2.3. *Collective incentives within PES (payment for environmental services) programs*

Collective payments in PES are a form of a group-based voluntary policy (Kotchen and Segerson, 2019; Segerson, 2013). 'Collective PES' can describe various arrangements, such as: groups must agree to establish a PES contract and supervise each other (Kerr et al., 2014), the conditions for payment require group performance (e.g., Kaczan et al., 2017; Salk et al., 2017), PES are applied within communal lands (e.g., Hayes et al., 2015; 2017), or whatever payments are made are received by the group instead of by the individual actor (e.g., Agrawal et al., 2015; Clements et al., 2010).

The effects of collective PES on cooperation, motivations and environmental performance have been studied both in the field, i.e., within actual PES programs, and within experimental settings. Looking at actual programs implemented in the developing world, Agrawal et al. (2015) find that participants in northern India who had engaged in communal conservation activities and received collective livelihood benefits from communal assets were less likely to shift from environmental to economic motivations for forest conservation, relative to those who receive private economic benefits. Consistent with that impact on orientation, Hayes et al. (2015) explore how a program in communal lands in Ecuador is related to the development of rules to manage common lands. They find that most communities have strengthened land-use rules since the implementation of a PES program. In a later study, Hayes et al. (2017) find PES reduced the number of households grazing livestock in collective land. They conclude that participation in PES programs reinforces communal resource-management arrangements. They also find collective PES might be more effective in groups with more of a history of collective institutions for management of resources.

Clements et al. (2010) compared PES programs in Cambodia which vary in the type of payment, i.e., whether it is made to individuals or villages. They find that the collective PES contracts are more institutionally effective and more widely supported locally than individual-based contracts. Sommerville et al. (2010) examine community-based programs in Madagascar, finding that the participants expressed perceptions of fairness within and, except for high opportunity costs, also net benefits from the PES programs. For some, poor governance diminished perceived benefits. Finally, Narloch et al. (2017) study pilot programs in the Bolivian and Peruvian Andes and find that collective PES may in fact yield better conservation incentives than individual payments. They also find that collective payments could mitigate some possible rent-seeking behaviors.

Experimental studies have produced a mix of results. Some find that collective payments to be ineffective and even crowd-out social norms by encouraging free riding (Narloch et al., 2012 for the Bolivian and Peruvian Andes; Midler et al., 2015). Moros et al. (2007), in contrast, find that responding to collective incentive payments enhances social motivations to protect forests. Using collective PES contracts with collective penalties for non-compliance, Kaczan et al. (2017) find, if anything, crowding-in of contributions by collective PES. Finally, Salk et al. (2017) find that group payments resulted in more significant short-term reductions in resource uses, compared with individual payments, due to communication and coordination and the perception of fairness.

Thus, as pointed out by Kotchen and Segerson (2019), the performance of group-based policies depends on designs and group interactions. Studies about collective PES highlight both points: the effect of a given policy based on environmental outcomes depends on local institutions; and the details of the policy can affect local trust, social capital, and motivations. Thus, there can be 'virtuous' or 'vicious' interactions between policies and the effective local institutions: a design may enhance or hinder cooperation, given local incentives to free-ride; and, given that, trust and coordination building are affected by the designs. In line with this, our experiment explores the effect of policy stringency on the potential for coordination through learning and trust building.

3. Setting

Small-scale gold mining is often *de facto* open access, lacking strong property rights (Shoko, 2002). That is the case in Colombia. The state owns the subsoil, and grants concessions,[8] yet according to the Colombian Mining Census, 86% of metallic mineral production occurs in small production units without mining titles (Cabrera and Fierro 2013). Most of those are

[8] In Colombia, there are approximately 9400 mining concessions (43% for gold), which cover 5.6 million hectares.

informal, i.e., lack permits and other requirements such as environmental licenses (Sarmiento et al., 2013). It is said that Colombia has about 180,000 small-scale gold miners (Cremers and de Theije, 2013) and, further, that they generate the highest per-capita mercury emissions in the world (Siegel, 2013).

About 40% of gold production in Colombia is in the Pacific region inhabited mainly by Afro-Colombian communities with collective land titles.[9] They are managed by community councils – collective organizations required by law to request titles – with locally elected representatives who are responsible for management of natural resources (Velez, 2011)[10]. Over 90% of the gold produced in the Pacific region comes from small-scale informal mining by locals and migrants, the latter often connected with armed groups (Sarmiento et al., 2013, Giraldo and Muñoz, 2012).

In the Pacific region, SSGM is a culturally embraced and traditional economic activity practiced since colonial times (Sarmiento et al., 2013). While in the past most mining was carried out using artisanal tools such as pans, since the superficial gold has been depleted, traditional tools have been combined with machines such as pumps, for small miners, or far bigger machines such as backhoes and dredges which help to move land and other materials in the search for gold.[11] Pumps are used to take water from mines or to separate material containing gold from riverbanks by using water pressure to flush out gold deposits. Their impacts on forests, sedimentation, and the paths of rivers increase with their number. In Colombian law, the only 'artisanal' mining – not requiring concessions – is via pans (*barequeo*). Thus, small miners face the same regulations as larger-scale mechanized miners. Yet communities consider low-horsepower pumping to be 'artisanal'. Community councils craft SSGM rules on, for example, revenue sharing and allowed mining techniques. They are also the local actors who could try policy with collective incentives.

We worked with mining communities in the rural areas of the municipality of Buenaventura, Colombia's main Pacific port. Despite the port's economic importance, poverty is rife. About 80% of people are below the poverty line and SSGM is the only economic alternative for many. Communities are organized in councils on the coast, along rivers, and along roads that connect to inner cities. River and road communities have significant mining. We focus on the seven mining communities located nearer to roads and, therefore, more closely connected to markets in cities.

Most of the communities we visited have migrant miners, who use heavy machines and mercury and compete for resources with local, less mechanized miners. Given the challenge of exclusion, encroachment by migrants is a feature of alluvial mining. Formal rights for subsoil resources contrast with *de facto* open access to alluvial deposits, though external miners often negotiate with councils to enter into their territories. As we were told in interviews, in some agreements councils required a share of the gold extracted to be paid as compensation to the community, with a portion of that going to the council and a portion going to landholders.[12] Such agreements are certainly not formally 'legal', as the Colombian legal framework very clearly establishes that all mineral resources belong to the state. However, given the states' incapacity to regulate SSGM on the frontier, community councils play *de facto* regulatory roles in controlling mining impact – albeit with mixed success, since power can be asymmetric in confrontations with external actors.

Given the significant differences between types of miners, and the existing security concerns, only local miners who are traditional members of Afro-Colombian communities with collective land titles participated in our 'field laboratory' experiments on collective incentives. Out of about four thousand total inhabitants within these communities, 640 adults participated in our experiments.

4. Experimental design & implementation

4.1. Design

We framed individual miners' choices in terms of pumps used – a measure of intensity distinct from mercury (or other pollution), as community members expressly oppose the use of mercury. Pumps serve as a metaphor for mining intensity, understood as a proxy for sedimentation, land degradation and deforestation. Such impacts are directly proportional to the number of pumps.[13] Our experiment refers to external damage, a downstream externality such as of sedimentation.[14]

Participants played in groups of five miners, each deciding on a number of pumps, i.e., choosing between 0, 1, 2 or 3 pumps. The collective target for a group is a number of pumps for the group. We varied this number according to the treatment: specifically, total targets of 0, 5 or 10 pumps. If the group's total pumps are at or below this maximum number, a collective reward is granted to the group and distributed equally among group members, regardless of the members' choices.

[9] Collective land titles do not exist everywhere in Colombia (but high monitoring costs can drive a collective approach).

[10] To date, almost 6 million ha have been collectively granted to more than 170 communities in the Pacific region.

[11] Techniques vary widely. Mercury is used by migrant miners but some Afro-Colombian and other communities have social norms against its use (Sarmiento et al., 2013 and our interviews, focus groups, and surveys in 2013/14).

[12] Collectively titled communities allow *de facto* forms of individual land holding based on kinship relationships.

[13] This brings us back to non-point pollution, where input taxes are an alternative to ambient taxes when the inputs are individually observable and are a dependable indicator of emissions. Pumps are an input that is observable but the cost for the state to observe them and determine who owns them is enormous, within such distant communities.

[14] Putting this in terms of Giordana and Willinger (2013)'s survey of experiments on non-point pollution, there is no internal damage to group members from the negative externalities (e.g., Cochard et al. 2005). Damage is purely external, implying no interactions among these firms other than those created by the incentive (e.g., Spraggon, 2002).

Table 1
Payoffs.

My decision (number of pumps) Mark with an X		Earnings from mining	Collective reward	Total earnings
0	0	$1	+$7	$8
1		$3	+$7	$10
2		$5	+$7	$12
3		$7	+$7	$14

Table 1 shows the payoffs, as they were shown to participants. The first column shows mining earnings for any number of pumps used. The second column shows the individuals' gain from the collective reward if the group complies with the target. The third column adds the earnings across the first two columns to show the individuals' earnings if the group does meet the target.

In each period, participants marked their decisions on a sheet containing Fig. 1, marking the number of pumps they chose with an X. A facilitator collected all the sheets, added up the total number of pumps, and communicated that total and individual earnings to each participant on the same sheet. Participants did not have previous knowledge of the number of rounds to be played, nor that the target might shift. We ran sessions with four or five groups, avoiding contamination among the groups by communicating totals in writing, using decision sheets (as just explained). Each facilitator was responsible for one group. Participants knew who were the other members of their group. At session's end, participants did a survey on their socioeconomic characteristics.

This type of coordination game features multiple symmetric and asymmetric Nash equilibria. A first pure-strategy, symmetric Nash equilibrium is zero contributions for all participants, which means that all individuals choose the maximum number of pumps. In the literature on threshold-public-good games, this is known as the ‘strong free-riding’ equilibrium (see Cadsby and Maynes, 1999). The second pure-strategy symmetric Nash equilibrium involves equal decisions by all that sum up to the threshold, i.e., the ‘threshold symmetric’ equilibrium. Unlike strong free-riding, it is efficient. Symmetries may create focal points (Cadsby and Maynes, 1999). In our game, a target of zero pumps has only the two symmetric Nash equilibria – i.e., zero or three pumps per person – while targets 10 or 5 have ‘strong free riding’ and ‘threshold symmetric’ as well as a number of asymmetric pure-strategy Nash equilibria. The latter are threshold equilibria, with total pumps at the group target, which maximizes the group's joint profits (given that both under- and over-compliance are inefficient).[15] However, the number of pumps and thus the contributions differ.

Individual incentives are affected by others' choices. If all others in a group choose a behavior that is one of the Nash equilibrium actions, then a player's best response is to choose that action, for inefficient or efficient equilibria. However, one deviation could imply that it is in everyone's interest to deviate. As the target decreases, the risks from others' possible deviations increase, in light of greater costs of compliance for each individual but the same benefit from compliance.

[15] For linear payoffs, an efficient Nash equilibrium in a threshold-public-goods game is any vector of individual contributions that: sums to the contribution threshold (efficiency constraint); and does not involve any individual contributing more than her benefits from the public good (rationality constraint) (Croson and Marks 1998, 2000).

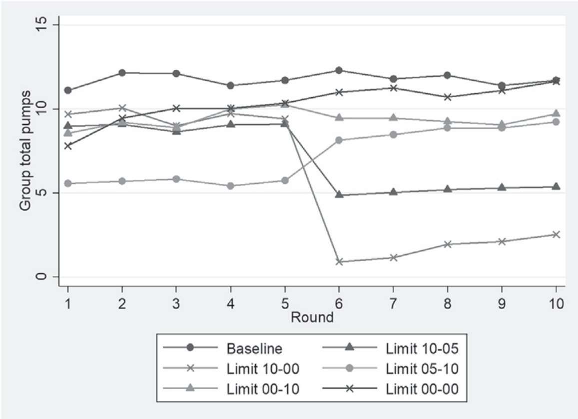


Fig. 1. Means per treatment of group totals.

Consequently, as the target decreases the efficient equilibria should become more unstable as even a 'small' deviation by one participant can move all others towards the 3-pump or inefficient equilibrium. For a zero target in particular, even a one-pump deviation by a single member ruins the group's effort, as no other member can counter this effect through her own choice of pumps.

In Table 2, for any total target (the rows), we show best responses, for any individual member, to any expectation they may hold about the number of total pumps chosen by others (the columns). The cells highlighted in blue are the symmetric and efficient Nash equilibria. The cells in orange are the asymmetric and efficient Nash equilibria. Finally, cells in green are inefficient equilibria.

4.2. Treatments

We explore collective rewards by varying the environmental targets, and their sequencing, across treatment groups. We want to study the potential for these communities to solve the coordination problems that such group-performance instruments involve. Some types of collective incentives, such as lump-sum group fines or rewards that we use, face coordination problems due to multiple Nash equilibria. Some kind of collective action, or coordination, is required to make them work. We hypothesize that coordination within groups rises if a group succeeds in meeting a target. Failures, however (including when they are likely due to more difficult targets, all else equal), could hinder the building of bonds and thus also lower the coordination to achieve future targets.

To test such dynamics, we randomly selected groups for sessions that employed different targets, which could vary over time − increasing or decreasing in stringency (vs. remaining constant) – using a between-subjects design with six treatments: a no-target baseline and five treatments that employ increasing, decreasing or constant stringency of environmental targets. In each treatment, group members made decisions for ten rounds in total. From the sixth round onward, though, sometimes we changed the target, i.e., the required contributions threshold, as is specified below:

- five rounds of target = 10 followed by five rounds of target = 05 (labeled "10 → 05")
- five rounds of target = 10 followed by five rounds of target = 00 (labeled "10 → 00")
- five rounds of target = 05 followed by five rounds of target = 10 (labeled "05 → 10")
- five rounds of target = 00 followed by five rounds of target = 10 (labeled "00 → 10")
- five rounds of target = 00 followed by five rounds of target = 00 (labeled "00 → 00")

The last treatment explores how groups converge, or diverge, if facing the strictest total target. The target '00' was framed as exclusive use of panning, i.e., literally using no pumps. For some groups, we framed changes in the targets as being due to either the entrance or the departure of external miners. For the other groups, we framed changes in targets as a decision by the state agency. We did not find any effects that were caused by framing differences, so we pooled each treatment across the external-miner and the agency framings. Table 3 summarizes our treatments.

4.3. Participants

We recruited 640 inhabitants of seven mining communities in the rural area of Buenaventura, via open invitations to all adults in these communities, including community leaders. Each session included 20−25 participants (4 or 5 groups). We split the participants randomly across groups and allowed one member per household in each group. Table 4 shows

Table 2
Individual best responses to others' behavior.

E(X-i)	0	1	2	3	4	5	6	7	8	9	10	11	12
$\bar{X}=0$	0	3	3	3	3	3	3	3	3	3	3	3	3
$\bar{X}=5$	3	3	3	2	1	0	3	3	3	3	3	3	3
$\bar{X}=10$	3	3	3	3	3	3	3	3	2	1	0	3	3

Rows are targets. Columns are expected totals of others pumps. Cells are best responses to those targets and those expectations (blue = efficient & symmetric Nash equilibria, orange = efficient & asymmetric Nash equilibria, green = inefficient equilibria).

Table 3
Treatments and number of participants.

	Treatments: Ordering of Stringency in Aggregate Pump Targets					No Policy
	10 → 05	10 → 00	05 → 10	00 → 10	00 → 00	
Framing with change in target as cause by externals miners entrance or departure	60 people (12 groups)	40 people (8 groups)	55 people (11 groups)	–	–	–
Framing with change in target as cause by government decision (no external miners)	120 people (24 groups)	55 people (11 groups)	60 people (12 groups)	100 people (20 groups)	100 people (20 groups)	50 people (10 groups)

participants' characteristics: 60% were women; average age was 36 years; and 68% had participated regularly in voluntary

Table 4
Socio-demographic characteristics road communities.

	Pooled	by treatment					
		none	10 → 05	10 → 00	05 → 10	00 → 10	00 → 00
Age (years)	35.7	36.4	36.0	31.3	36.4	38.2	35.9
Women (%)	60.1	54.0	63.9	62.8	63.5	60.0	50.0
Voluntary participation (%)	67.9	66.0	66.1	53.2	66.9	73.0	81.8
Education (%)							
none	6.9	10.0	8.3	5.3	7.0	5.0	6.0
Primary Incomplete	21.4	22.0	25.6	14.9	20.0	20.0	23.0
Primary Complete	12.4	14.0	11.7	7.4	16.5	14.0	11.0
Secondary Incomplete	25.2	18.0	23.9	29.8	24.3	31.0	22.0
Secondary Complete	23.9	20.0	24.4	28.7	26.1	15.0	27.0
Technical or College	10.2	16.0	6.1	13.9	6.1	15.0	11.0
Artisanal/Small Miner (%)	75.0	76.0	76.1	80.0	71.3	77.0	70.0
Income, good week (US$)	153.4	115.7	153.0	212.1	116.3	161.9	149.8
Income, bad week (US$)	22.1	21.2	17.0	31.5	19.4	19.6	29.2

community activities. In terms of formal education: 7% had none, 21% did not complete primary school, 12% had completed primary school but no more, 25% had entered but did not complete secondary school, 24% completed secondary school, and 10% went beyond secondary education. The main work activity is artisanal or small-scale gold mining for 75% of participants. Their average reported income in a good week is about US$150, and in a bad week it is about US$20.[16]

Looking at the differences in the participants' sociodemographic variables across our treatments (that were randomly assigned but, nonetheless, could differ somewhat given small numbers): the baseline and the 00 → 00 sessions featured a lower percentage of women; the 10 → 00 participants happened to participate less in community activities; the 00 → 00 participants happened to have the highest rates of participation in community activities; and the 10 → 00 participants happened to have more years of formal education as well as higher earnings than did the other participants. We controlled for these differences using regressions that are presented in the following section.

[16] To select communities, as well as the framing for the experiment, we first characterized the mining communities near Buenaventura using secondary information. Next, we interviewed the leaders of those communities as well as local authorities to elucidate relevant characteristics of those communities and the features of small-scale mining.

Table 5
Mean total pumps & compliance rates, across groups, by treatment and by stage.

	1st Stage Target	First Stage (rounds 1–5)		2nd Stage Target	Second Stage (rounds 6–10)	
		Mean Pumps	Compliance Rate		Mean Pumps	Compliance Rate
No Targets (baseline)	**–**	11.74 (2.32)	–	–	11.84 (2.61)	–
(1) 00 → 00	**00**	9.54 (5.14)	15.0% (0.36)	**00**	11.14** (5.0)	14.0% (0.35)
(2) 00 → 10	**00**	9.38 (5.47)	20.0% (0.4)	**10**	9.38 (2.48)	76.0% (0.43)
(3) 05 → 10	**05**	5.65 (1.98)	65.2% (0.48)	**10**	8.71*** (1.41)	92.2% (0.27)
(4) 10 → 05	**10**	8.97 (1.75)	86.1% (0.35)	**05**	5.15*** (1.4)	66.7% (0.47)
(5) 10 → 00	**10**	9.58 (1.45)	75.8% (0.43)	**00**	1.73*** (2.99)	55.8% (0.5)

Standard deviations in parentheses ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ for comparison of mean group total pumps between stages.

5. Results

5.1. Legacy building & average treatment effects

While our core question is whether collective rewards can induce groups to use fewer pumps, the corollary question is whether they achieve stated targets. Table 5 shows average pumps – seen by round in Fig. 1 – and success in compliance by treatment and by stage. In our baseline, groups averaged about 12 pumps, above all targets but significantly lower than the 15 maximum. The maximum is predicted, lacking any collective reward, since the largest individual earnings arise when each group member has chosen three pumps. Baseline choices lower than 15 pumps may be due to motivations of participants other than economic earnings. For example, we stated in the instructions that more pumps imply greater earnings but also more environmental damage. Relative to the baselines, total pumps were always lower with targets and collective incentives (except for late rounds in the toughest possible sequence 00 → 00, when they are about the same).

5.1.1. Challenging target (00): short-run failure, negative legacy, yet efficient and symmetric

Table 5's rows (1) and (2) convey the overall failures of these groups to overcome challenges posed by the most stringent targets of zero pumps. Even if successful, this treatment has payoffs ($8 each) only slightly above those ($7 each) from the inefficient Nash of three pumps each. Further, success can be ruined by a single one-pump deviation. Perhaps it is not surprising, then, that compliance is under 20% in first stages and total pumps are far above zero, in fact over nine.

Table 5 also shows that groups that faced a zero target during the first stage carried forward a negative legacy of lack of coordination, due to prior failures. First, if the target remains at zero, in (1), the total for pumps rises significantly, effectively achieving a total treatment failure, since the pump total rises to become basically equal to the baseline. Second, comparing (2) and (3), both of which have the easiest target of 10 pumps in the second stage, we see that the successes are considerably higher for 10 given more successes for 05 than for 00. Further suggesting such a negative legacy of prior failure, most of those differences occurred early in 2nd stages (Fig. 1).

Table 6 highlights two issues that could matter from the perspective of these miners' preferences. First, if compliant, miners do not want to be overcompliant but at the efficient Nash equilibrium equal to the target: once a group reward is achieved, any overcompliance lowers their earnings.[17] Second, some may prefer that any efficient Nash equilibrium be symmetric, i.e., the same for all.

Recall that for zero pumps, compliance is *always* both efficient and symmetric (rows 1 & 2). There is only one efficient Nash equilibrium: all zeros. Yet the rate of compliance for zero is low compared to the other treatments. Thus, while compliance is always efficient, it is also the case that non-compliance is high. Overall, then, efficiency is lower than for the less stringent targets.

5.1.2. Medium target (05): short-run success, yet less often efficient or symmetric

Intermediate stringency of the environmental target, i.e., five (05), is seen in Table 5's row (3), for 05 in the 1st stages, and row (4), for 05 in the 2nd stages. These rows communicate a fair bit of success and thus impact. About two thirds of groups manage to comply, obtaining the collective reward. Some do not. Yet, still the total numbers of pumps are under six on average, i.e., almost down to 05. These outcomes are also clear in Fig. 1, while the dynamics by round are clearer.

Further, as noted, if a target of 10 followed a target of 05 (row 3), then almost 92% of the groups achieved compliance, not suggesting any negative legacy if comparing to 10 in rows (4) or (5). If anything, there may have been a positive legacy of coordination, from managing to achieve 05: in rows (4) and (5) the initial target was the easier 10 and the average compliance was only 80%.

However, Table 6 shows that options for groups in ways to achieve five yields overcompliance (total pumps below target) and asymmetric compliance (some participants have more pumps). For rows (3) and (4), about half of compliant cases (the

[17] If payments depend on distance between pollution and the threshold, overcompliance can be beneficial to firms leading to collusion (e.g., Suter et al., 2008). With fixed fines or rewards, there are no such gains in overcomplying.

Table 6
Compliance across groups, by treatment and by stage.

	1st Stage Target	First Stage (rounds 1–5)			2nd Stage Target	Second Stage (rounds 6–10)		
		Compliance	Efficient Compliance	Symm. Effic. Compliance		Compliance	Efficient Compliance	Symm. Effic. Compliance
(1) 00 → 00	**00**	15.0%	15.0%	15.0%	**00**	14.0%	14.0%	14.0%
(2) 00 → 10	**00**	20.0%	20.0%	20.0%	**10**	76.0%	21.7%	6.0%
(3) 05 → 10	**05**	65.2%	39.1%	28.7%	**10**	92.2%	30.5%	3.5%
(4) 10 → 05	**10**	86.1%	30.6%	10.0%	**05**	66.7%	38.3%	15.6%
(5) 10 → 00	**10**	75.8%	29.5%	12.6%	**00**	55.8%	55.8%	55.8%

~66% seen in Table 5) are efficiently achieving a five target. Thus, about as many cases involved some overcompliance. Further, only half of the efficient cases are symmetric (varying by row). At least some members overcomply; i.e., choose no pumps, perhaps reflecting worries that others may not select the efficient 1 pump.

5.1.3. *Easy target (10): short-run success, positive legacy, yet less often efficient or symmetric*

Table 5's rows (2–5) show that groups solved the coordination challenge posed by a target of 10. Averaging across the 2nd stages in rows (2) and (3), as well as the 1st stages in rows (4) and (5), we can see that over 80% of groups managed to comply in terms of total pumps. Even average pumps 'are compliant' at about 9, for the 2nd stages, or just above 9 for the 1st stages. This does not imply much impact in relation to the baseline but, still, clearly does imply some coordination.

Despite being small, the impacts from succeeding for this easiest target had a positive legacy. Table 5 finds that groups facing 10 did as well or far better later on. When the 2nd target was 05, in row (4), total pumps are lower and compliance higher that for target 05 in the 1st stages. Even clearer are the choices for 00: row (5) finds, in contrast to under 20% compliance in (1) and (2), that facing 00 in 2nd stages, compliance was over 55% after 1st-stage success for 10. In Fig. 1, coordination falls yet is far better than in 00-00. Half the groups fail yet average pumps fell to 2.

Table 6 shows again that flexibility – here in ways that a group can meet a target of 10 – lowers efficiency and symmetry if compliant. Around one third of compliant groups are efficient for 10. Further, on average, only about one third of the cases of efficient compliance are also symmetric. This raises dynamics. If compliance is inefficient, it can lead to adjustments towards efficiency, which can lead to non-compliance that can undermine a positive legacy. The same can occur for asymmetry, i.e., inequity within efficient compliance, although Fig. 2 indicates some stability.

5.1.4. *Average treatment effects*

Table 7 examines compliance in a probit regression, clustered by group, with robust standard errors and a dummy for each target (logit confirms robustness). As the baseline has no target, the omitted treatment is 10 in the 2nd stage. As above, greater stringency yields lower compliance: 05 is worse than 10, 00 is worse than 05, worse still if 00 is first, and worst of all if 00 continues.

Table 8 presents an OLS regression for total pumps, clustered by group with robust standard errors, with the baseline as the omitted treatment and a dummy for each target (for robustness we have a Poisson regression for total pumps in Appendix 1). Regressions confirm what the tables suggested. For example, in relation to the baseline, the average number of pumps per group falls by 2–3 pumps for a target of 10 and 6–7 pumps for a target of 5. Thus, across the relatively more feasible environmental targets (i.e., not 00), raising the stringency lowers the number of pumps.

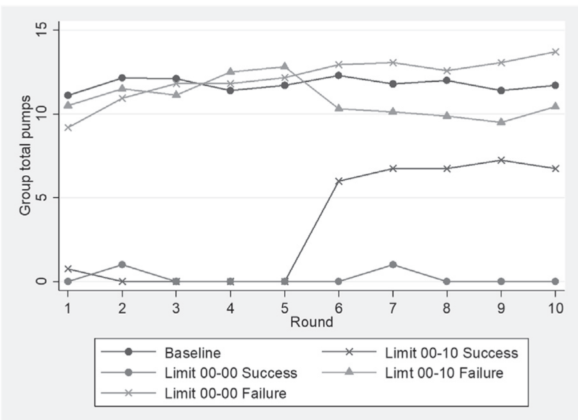


Fig. 2. Means for baseline and treatments starting with zero.

Table 7
Explaining groups' compliance rates.

Probit regressions	(1)	(2)	(3)
1 if target = 0 s	−1.340***	−1.006***	−1.069***
	(0.256)	(0.254)	(0.254)
1 if target = 5 s	−0.697***	−0.696***	−0.626***
	(0.226)	(0.227)	(0.228)
1 if target = 0 first	−0.563**	−0.575**	−0.570**
	(0.224)	(0.225)	(0.223)
1 if target = 5 first	−0.000600	−0.000875	−0.0548
	(0.251)	(0.251)	(0.254)
1 if target = 10 first	−0.115	−0.114	−0.0800
	(0.206)	(0.206)	(0.206)
1 if target = 0 first & second		−0.739*	−0.749**
		(0.381)	(0.381)
Women in the group (%)			−0.434
			(0.335)
Less than primary education (%)			−0.664
			(0.407)
Constant	1.345***	1.343***	1.746***
	(0.241)	(0.241)	(0.286)
Observations	1180	1180	1180

Robust standard errors (clustered by groups) in parentheses − Includes controls per session.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ (& omitted category is a target of 10 in the second stage).

Table 8 also confirms that a 00 target works quite differently and illustrates conditions in which it works particularly poorly.[18] In Table 8, model (5), the effect of a 00 target in the second stage is estimated to be a reduction in pumps of 7.7, compared to the baseline. Yet the effects of facing a zero target in the first stage and, worse still, facing zero in all rounds, cancel the target's gains.

5.2. Legacy building & group dynamics

5.2.1. Divergent paths facing the toughest challenge

Most groups facing the toughest target of zero, in initial rounds, were not able to comply despite it being the sole efficient equilibrium: three out of twenty groups in 00 → 00; and four out of twenty in 00 → 10 (the same treatment for early rounds, as participants did not expect a change). Table 9 and Fig. 2 separate initially compliant from initially non-compliant. In each treatment, we see that for 00, initial rounds reveal, or determine, the subsequent ability of groups to comply. Table 9's top rows show that first-stage mean total pumps are essentially equal to the baseline for initially non-compliant groups. For these groups, the points above are more severe. When the challenging target of 00 continues (row 1), a negative legacy of failure is seen in the 2nd stage, with 13 pumps, i.e., significantly larger compared to the 1st stage and a more inefficient outcome.

If the target 00 is followed by the weakest target 10 (row 2), there is a striking and significant reduction in the mean pumps per group, even though the target has gone from hardest to weakest. This can result from restoring small but clear expectations versus a collapse of ambitious hopes. However, a negative legacy remains to some extent, as that average is still above the target of 10, while the success rate is only 71%, contrasting with the high 2nd-stage success in rows 3 and 4. Thus, the initially compliant groups that achieved the tough target to earn the collective reward appear to generate a positive legacy of ongoing coordination despite a very small individual gain.

5.2.2. Searching for efficiency in coordinated compliance

Table 10 considers the dynamics over rounds, as groups react to prior successes and failures. A group member can never be sure whether her group will comply or whether her decisions affect compliance. In each round, members try to learn something about other members, via decisions. Groups above targets may adjust to comply. Those below targets may raise numbers of pumps.

To examine such dynamics, Table 10 explores the change in the number of pumps from one round to the next ($pumps_t - pumps_{t-1}$). The key independent variable for exploring the dynamics is distance from the target in the previous round (target minus total pumps, ≥ 0 when compliant). Consistent with efficient adjustment, its effect is positive. Thus, when inefficiently compliant, i.e., below the target as members were 'too responsible', a group's mean number of pumps rises. Yet if being

[18] It is worth noting one additional result: pumps chosen rise with the expectations one has for others' pump choices (see Appendix 2). That suggests a prevalence among our participants of 'conditional cooperation' (Chaudhuri, 2010).

Table 8
Explaining groups' total pumps.

OLS regressions	(1)	(2)	(3)	(4)	(5)
1 if target = 10 s	−2.678***	−2.463***	−2.603***	−2.629***	−2.617***
	(0.700)	(0.726)	(0.764)	(0.754)	(0.773)
1 if target = 5 s	−6.444***	−6.030***	−6.172***	−6.246***	−6.451***
	(0.705)	(0.739)	(0.739)	(0.722)	(0.751)
1 if target = 0 s	−3.765***	−3.924***	−5.416***	−7.750***	−7.601***
	(1.025)	(1.054)	(1.212)	(1.019)	(1.034)
1 if target = 10 first			0.259	0.243	0.157
			(0.323)	(0.319)	(0.318)
1 if target = 5 first			0.390	0.408	0.567
			(0.392)	(0.388)	(0.408)
1 if target = 0 first			2.940***	2.994***	2.947***
			(1.000)	(0.922)	(0.906)
1 if target = 0 first & second				4.606***	4.614***
				(1.281)	(1.263)
Women in the group (%)					1.176
					(0.828)
Less than primary education (%)					2.048*
					(1.215)
Constant	11.79***	10.65***	10.62***	10.80***	9.639***
	(0.685)	(0.872)	(0.873)	(0.818)	(0.952)
Observations	1280	1280	1280	1280	1280
R-squared	0.193	0.212	0.253	0.352	0.367

Robust standard errors (clustered by groups) in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ (& omitted category is the baseline without a target).

inefficiently non-compliant, e.g., just above a target, the group's total pumps falls (Appendix 3 presents assemblages, by treatment, of figures showing each round for each group).

Once we control for such adjustment, we see that for the severe (zero) target, the core dynamic is to increase pumps, as we would expect if members lose confidence in a group. The members' lack of confidence in their groups' abilities to coordinate efficiently at zero pumps affects the efficient-adjustment dynamic. Models (3) and (4) in Table 8 interact a zero target with the prior distance of the group's total of pumps from the target. Adding those two coefficients shows that the efficient adjustment is essentially not happening for the more severe (zero) target. Further, having separated that out, we can see stronger adjustment dynamics for other treatments (again see our Appendix 3, with figures by round for each group, for an illustration of this difference).

6. Discussion

When individual producers' behaviors are impossible or very costly to monitor, interventions based on aggregate measures, when paired with local collective action, could provide incentives for collective compliance. We proposed a rewards mechanism for mining groups, conditional on aggregate environmental compliance. Our focus is on rewards since artisanal gold mining occurs far out on frontiers, where there are high levels of poverty that complicate fine-based solutions. Instead of individual behaviors, a state would need to observe an aggregate measure, such as the total mercury content in a sample of river water, total deforested area within satellite data, or perhaps simply the numbers and maybe also the types of machines observed in any mining area.

We explored collective compliance and individual choice for group targets of varied stringency – with different stringency sequences, i.e., treatments that increase in stringency versus decrease. We did not provide any mechanisms to support the groups' coordination, instead outsourcing that completely to local collective action (and without even allowing group communication, though adding communication may be a useful extension for robustness concerning sequences).

We found that collective conditional rewards were able to induce efficient group compliance and even overcompliance, sometimes meeting even the most stringent targets. However, the more stringent targets could be counterproductive in the 'short' and 'long' term. Most groups starting with the most severe target were not able to comply. Further, a complete collapse of coordination can yield environmental outcomes that are even worse than in the baseline. These results echo threshold-public-goods results in which hard thresholds raise contributions but also the chance for a target not to be met (Ledyard, 1995; Suleiman and Rapoport, 1992; Cadsby and Maynes, 1999). However, they contrast with lower emissions levels for harder targets in Suter et al. (2010).

The key novelty in our results is the dynamic effects of changing the stringency level over time, with apparent learning by groups via successful experiences that build a capacity to coordinate. Concretely, initial success for an easier target seems to help with a later more challenging target. Expectations of other participants' behaviors play key roles and are more positive after success.

Table 9
Average total pumps & overall compliance rates, across groups. given an initial target of zero and splitting by initial compliance.

	# of groups	*First Stage (rounds 1–5)*		*Second Stage (rounds 6–10)*	
		Mean Pumps	Success Rate	Mean Pumps	Success Rate
00 → 00 *(initially non-compliant)*	17	11.19 (3.57)	1.2% (0.11)	13.07*** (2.05)	0.0% (0.0)
00 → 10 *(initially non-compliant)*	16	11.69 (3.24)	1.3% (0.11)	10.05*** (1.99)	71.3% (0.45)
***00** → **00** (initially compliant)*	3	0.2 (0.77)	93.3% (0.26)	0.2 (1.99)	93.3% (0.26)
***00** → **10** (initially compliant)*	4	0.15 (0.67)	95.0% (0.22)	6.7*** (2.52)	95.0% (0.22)

Standard deviations in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1 for comparison of mean group total pumps between stages.

Table 10
Regressions for group behavior: change in pumps between rounds.

OLS regressions	(1)	(2)	(3)	(4)
1 if target = 5 s	0.0609	0.155	0.588***	0.559***
	(0.103)	(0.121)	(0.196)	(0.196)
1 if target = 0 s	0.207	1.027***	0.929***	0.975***
	(0.125)	(0.281)	(0.251)	(0.260)
1 if target = 10 first	−0.0677	−0.125	−0.0301	−0.0427
	(0.0843)	(0.0991)	(0.169)	(0.169)
1 if target = 5 first	−0.0835	−0.0303	0.158	0.186
	(0.110)	(0.125)	(0.212)	(0.214)
1 if target = 0 first	0.0435	0.469**	0.230	0.233
	(0.158)	(0.232)	(0.186)	(0.184)
Distance from target t-1		0.131***	0.490***	0.495***
		(0.0246)	(0.0721)	(0.0724)
Distance from target t-1 * target = 0			−0.434***	−0.435***
			(0.0733)	(0.0737)
Women in the group (%)				0.184
				(0.229)
Less than primary education (%)				0.333
				(0.254)
Constant	0.0517	−0.132	−0.554***	−0.743***
	(0.0874)	(0.135)	(0.202)	(0.246)
Observations	974	904	904	904
R-squared	0.004	0.064	0.161	0.163

Robust standard errors (clustered by groups) in parentheses – and including controls per session.
***p < 0.01, **p < 0.05, *p < 0.1.

Our results are relevant for mining in Colombia and other countries with similar challenges, such as Peru and Brazil. In international agreements, mercury is the main concern regarding SSGM. For example, the Minamata Convention states that the only appropriate mercury target is zero, and stringent mandates have also been linked to limits on the use of machines in mining. Indeed, even a small machine such as a pump can trigger the application of policies designed for larger mechanized miners, yet such strong limitations have been ineffective, consistent with our results. Gradually increasing stringency, to build coordination, could make compliance more feasible.

Such an approach could be crafted to respect local heterogeneity and choice. A menu of options could offer lower or higher rewards for lower or higher stringency. We believe that actual mining authorities can implement policies such as those we studied. However, as seen within our results, for strict targets, benefits of collective incentives could be offset by financial costs of compliance alongside risk aversion concerning coordination as well as costs of monitoring and coordination.

Future research could consider various other types of policies, e.g., combining carrots and sticks, as proposed by Segerson and Wu (2006): if miners do not comply voluntarily with the limit, they would face credible mandatory punishment (we are not sure whether, in fact, external sanctions can be credible on these frontiers). A combination of internal monitoring and sanctioning across group members, which could facilitate coordination with the tough policies, could also be tested. As noted, exploring the impacts on results of within-group communication seems worthwhile.

Research could also address more applied practical questions, exploring not only these concepts but also the most effective parameters for implementation, such as the type and size of rewards. Rules for distributing rewards within groups may be critical, with payments to individuals versus the provision of local public goods from which all individuals benefit (perhaps equally, perhaps not). Collective in-kind payments may provide stronger conservation incentives in cases where collective action is robust (Narloch et al., 2017). In-kind payments are relevant here, as miners in the field have expressed that local development projects (including productive activities, roads, or schools), or even regional permits for mining, could be desirable forms of collective rewards.

Uncertainty, as is inherent for any aggregated measures employed by a state, is likely to matter. Socially speaking, unpredictable arrivals of migrant miners in a region seem critical in Colombia. Mining damage is perceived by the mining communities themselves and also can be important, e.g., mercury is a stock pollutant that might affect miners as it accumulates in the environment. However, we suspect that there is a lack of sufficient information on these potential damages and more generally that there exist uncertainties regarding the effects of mercury on human health.

In sum, our results suggest that conditional collective incentives have the potential to allow for production by small-scale miners, which can generate a significant portion of their livelihoods, and yet also improve their health as well as their regions' environmental performance. Collective incentives or rewards for compliance could take various forms such as mining permits, access to certified gold markets, payments for ecosystem services, and public goods. Successes, however, will depend on the abilities of groups to build effective local strategies to coordinate decisions.

Acknowledgements

This study was conducted with support from Latin America and the Caribbean Environmental Economics Program (LACEEP) as well as the Center for Research on Environmental Decision (CRED) at Columbia University. We thank LACEEP's Scientific Committee for their useful comments and contributions. We also thank our field team for its assistance in implementing the experiments: Alvaro Arroyo, Lina Moros, Sergio Puerto, Maria Camila Hernandez, Valentina Fonseca, Pablo Ramos, Giovanni Laiton, Nicolas Gonzales and Deisy Gutierrez. Many thanks to the Buenaventura community councils' leaders and inhabitants for their support and participation.

Appendix 1. Explaining groups' total pumps (here Poisson version of Table 8)

Poisson regressions	(1)	(2)	(3)	(4)	(5)
1 if target = 10 s	−0.258***	−0.232***	−0.249***	−0.252***	−0.252***
	(0.0602)	(0.0645)	(0.0695)	(0.0677)	(0.0714)
1 if target = 5 s	−0.791***	−0.740***	−0.771***	−0.783***	−0.812***
	(0.0659)	(0.0716)	(0.0704)	(0.0675)	(0.0717)
1 if target = 0 s	−0.385***	−0.406***	−0.611***	−0.955***	−0.941***
	(0.111)	(0.116)	(0.161)	(0.159)	(0.160)
1 if target = 10 first			0.0301	0.0275	0.0150
			(0.0363)	(0.0355)	(0.0358)
1 if target = 5 first			0.0790	0.0819	0.104
			(0.0688)	(0.0686)	(0.0701)
1 if target = 0 first			0.371***	0.377***	0.373***
			(0.141)	(0.129)	(0.127)
1 if target = 0 first & second				0.593***	0.599***
				(0.171)	(0.168)
Women in the group (%)					0.167
					(0.104)
Less than primary education (%)					0.262*
					(0.145)
Constant	2.467***	2.327***	2.324***	2.352***	2.192***
	(0.0580)	(0.0922)	(0.0920)	(0.0826)	(0.107)
Observations	1280	1280	1280	1280	1280

Robust standard errors (clustered by groups) in parentheses.
***p < 0.01, **p < 0.05, *p < 0.1 (& omitted category is the baseline without a target).

Appendix 2. Initial individual pumps decision (Tobit & Poisson versions below)

Tobit regressions	Pooled		Baseline	Limit = 0	Limit = 5	Limit = 10
1if limit = 10	−0.703**	−0.675***				
	(0.288)	(0.249)				
1if limit = 5	−1.595***	−0.576**				
	(0.303)	(0.262)				
1if limit = 0	−1.099***	−0.316				
	(0.348)	(0.283)				
Expectations total pumps		0.273***	0.270***	1.265***	0.117**	0.108***
		(0.0165)	(0.0643)	(0.268)	(0.0474)	(0.0185)
Age (years)		−0.0101**	−0.00110	−0.156***	0.00690	−0.00145
		(0.00444)	(0.0162)	(0.0568)	(0.0066)	(0.00443)
Gender (1 if female)		0.182	0.876	0.829	0.361*	0.160
		(0.113)	(0.620)	(1.137)	(0.192)	(0.0989)
Less than primary education		0.332**	0.493	2.382	−0.271	0.210
		(0.147)	(0.588)	(1.520)	(0.219)	(0.131)
1 if voluntary work		−0.199*	0.114	−1.046	−0.0328	−0.166
		(0.111)	(0.615)	(1.238)	(0.154)	(0.102)
1 if mining primary activity		−0.124	−0.289	−0.0441	−0.0607	−0.146
		(0.108)	(0.489)	(1.169)	(0.160)	(0.100)
Constant	2.300***	0.631*	−0.236	−0.144	0.0979	1.022***
	(0.313)	(0.348)	(1.137)	(2.308)	(0.423)	(0.237)
Sigma	1.665***	1.213***	1.343***	4.946***	0.811***	0.779***
	(0.0831)	(0.0626)	(0.265)	(0.968)	(0.0974)	(0.0517)
Observations	640	624	46	189	115	274
Lower censored observations	110	108	4	80	15	9
Upper censored observations	171	162	24	88	8	42

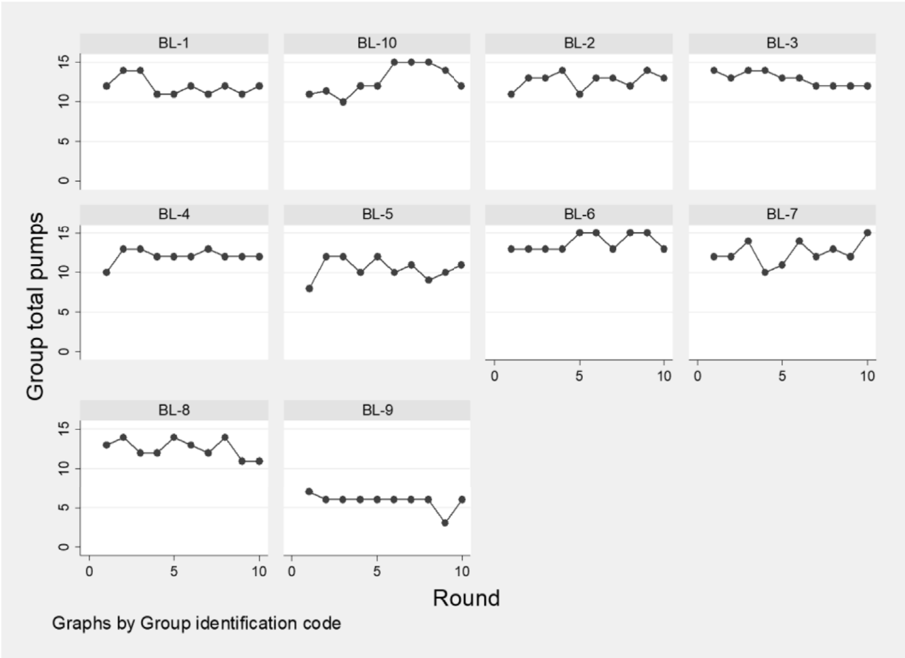
Robust standard errors in parentheses – Includes controls per session.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Poisson regression	Pooled		Baseline	Limit = 0	Limit = 5	Limit = 10
1if limit = 10	−0.143**	−0.0152				
	(0.0665)	(0.0645)				
1if limit = 5	−0.633***	−0.177*				
	(0.0870)	(0.0948)				
1if limit = 0	−0.323***	−0.0822				
	(0.0880)	(0.0708)				
Expectations total pumps		0.0897***	0.0603***	0.109***	0.0844***	0.0527***
		(0.00634)	(0.0163)	(0.0102)	(0.0319)	(0.00911)
Age (years)		−0.00461***	−0.00323	−0.0151***	0.00478	−0.00121
		(0.00161)	(0.00275)	(0.00429)	(0.00439)	(0.00200)
Gender (1 if female)		0.0781*	0.135	0.124	0.239	0.0655
		(0.0423)	(0.137)	(0.107)	(0.152)	(0.0463)
Less than primary education		0.133**	0.108	0.349**	−0.176	0.0925
		(0.0536)	(0.115)	(0.154)	(0.160)	(0.0581)
1 if voluntary work		−0.0672	0.0327	−0.0547	−0.0359	−0.0756*
		(0.0416)	(0.109)	(0.131)	(0.108)	(0.0456)
1 if mining primary activity		−0.0681*	−0.0366	−0.0770	−0.0179	−0.0697
		(0.0400)	(0.0860)	(0.105)	(0.118)	(0.0446)
Constant	0.626***	−0.00767	0.250	0.0536	−0.613**	0.195*
	(0.0815)	(0.122)	(0.278)	(0.250)	(0.267)	(0.109)
Observations	640	624	46	189	115	274

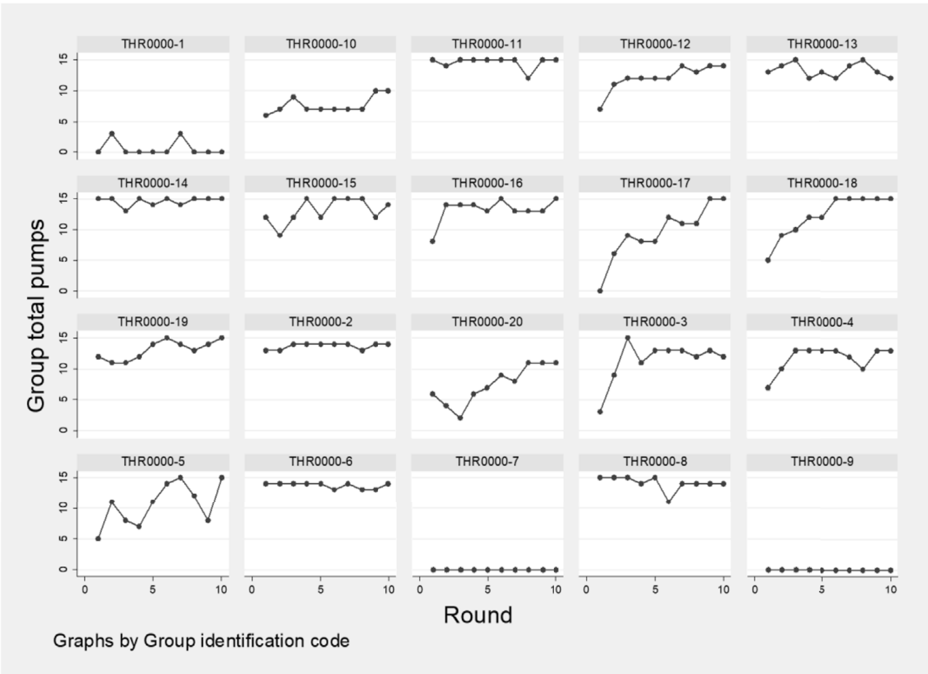
Robust standard errors in parentheses – Includes controls per session.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Appendix 3. Total pumps by round for each group, organized by treatment

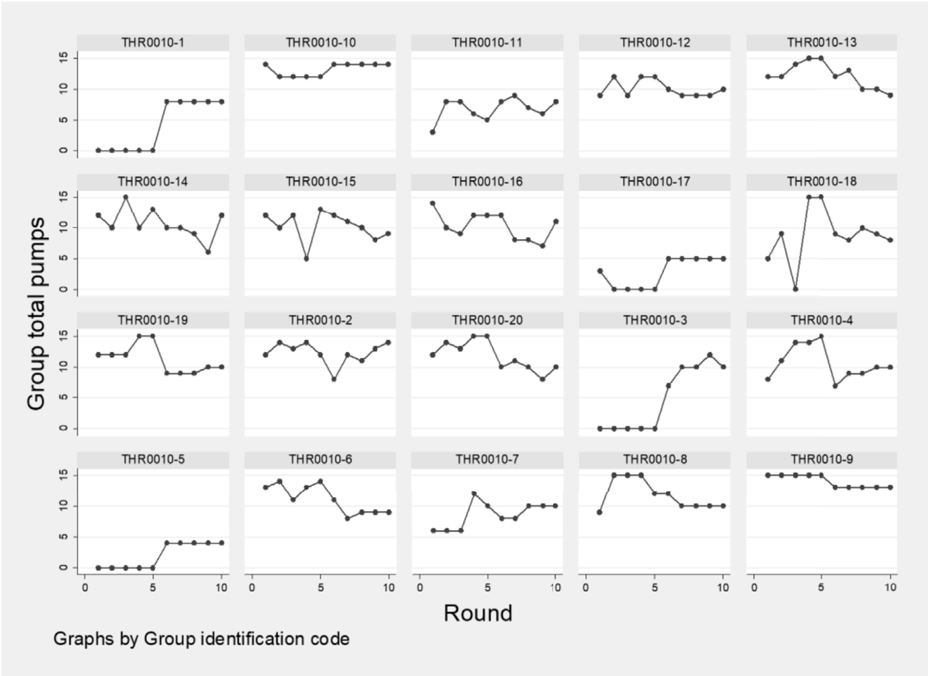
Baseline (no policy)



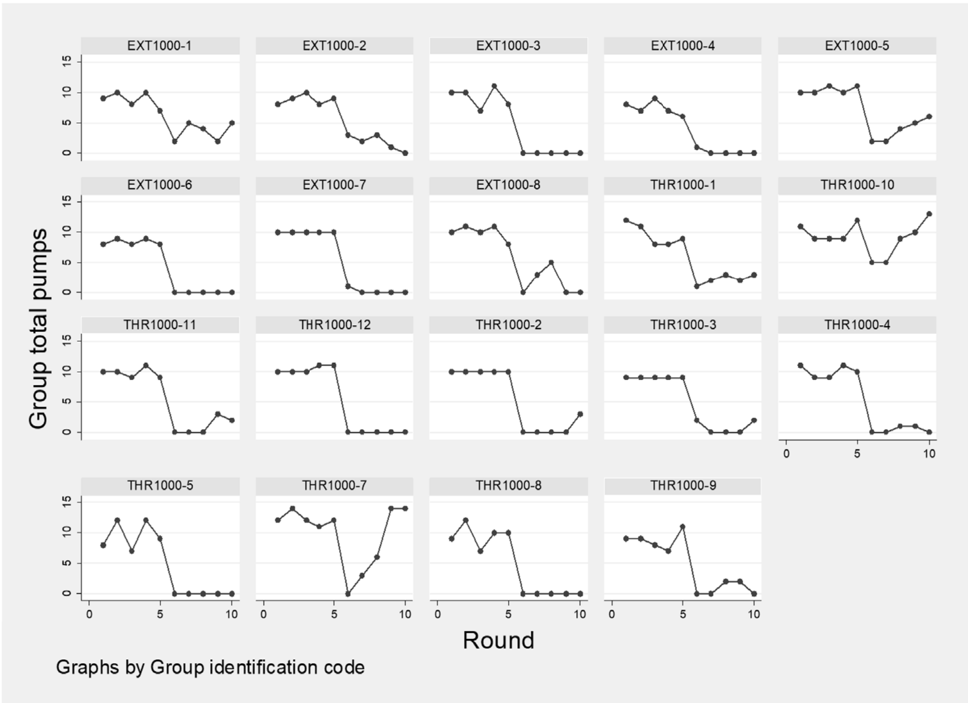
00-00 (i.e., zero target in first phase (Rounds 1-5), zero target in second phase (R6-10))



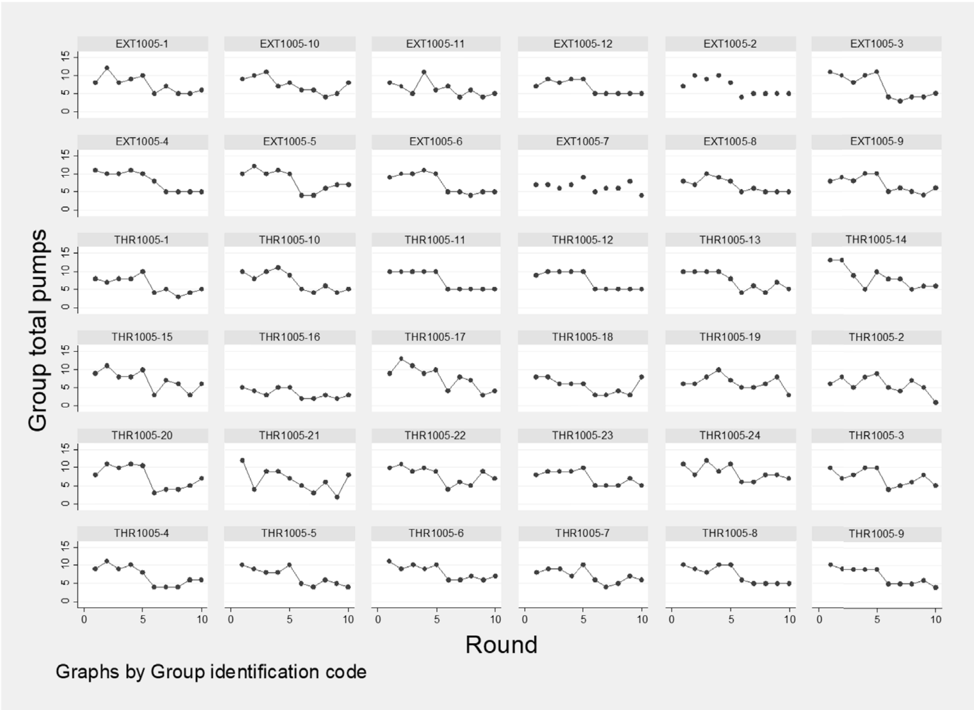
00-10 (i.e., zero target in first phase (Rounds 1-5), ten target in second phase (R6-10))



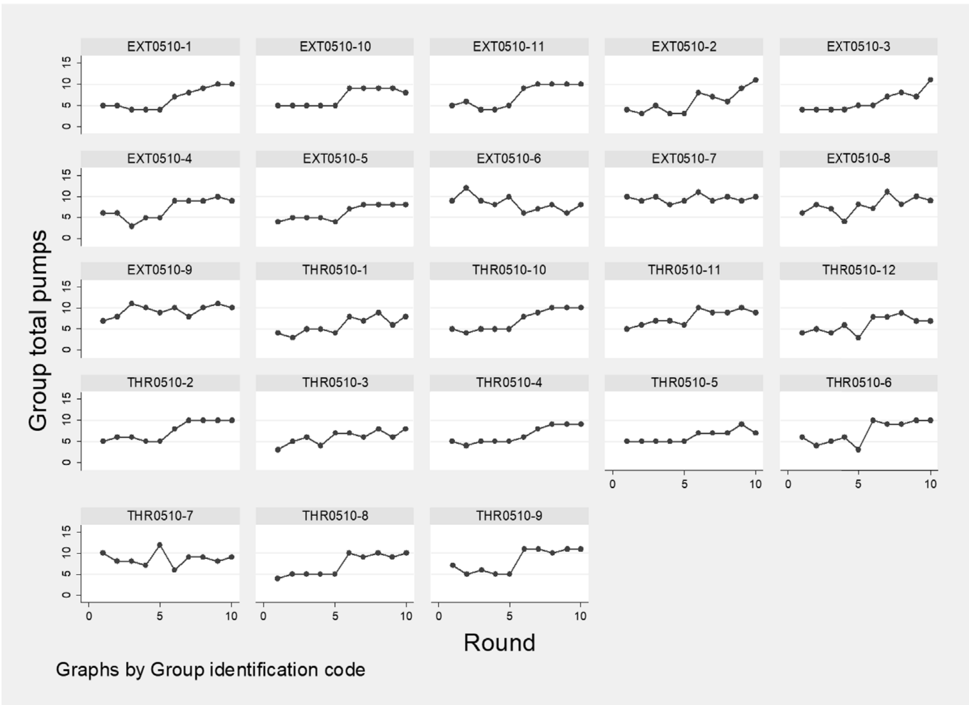
10-00 (i.e., ten target in first phase (Rounds 1-5), zero target in second phase (R6-10))



10-05 (i.e., ten target in first phase (Rounds 1-5), five target in second phase (R6-10))



05-10 (i.e., give target in first phase (Rounds 1-5), ten target in second phase (R6-10))



References

Agrawal, A., Chhatre, A., Gerber, E.R., 2015. Motivational crowding in sustainable development interventions. Am. Pol. Sci. Rev. 109 (3), 470–487.

Alpízar, F., Requate, T., Schram, A., 2004. Collective *versus* random fining: an experimental study on controlling ambient pollution. Environ. Resour. Econ. 29, 231–252.

Armendáriz de Aghion, B., Morduch, J., 2005. The Economics of Microfinance. MIT Press.

Asner, G.P., Llactayo, W., Tupayachi, R., Ráez Luna, E., 2013. Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. Proc. Natl. Acad. Sci. 110 (46).

Bagnoli, M., McKee, M., 1991. Voluntary contribution games: efficient private provision of public goods. Econ. Inq. 29, 351–366.

Barrett, S., Dannenberg, A., 2012. Climate negotiations under scientific uncertainty. Proc. Natl. Acad. Sci. 109 (43), 17372–17376.

Barrett, S., Dannenberg, A., 2017. Tipping versus cooperating to supply a public good. J. Eur. Econ. Assoc. 15 (4), 910–941.

Besley, T., Coate, S., Loury, G., 1993. The economics of rotating savings and credit association. Am. Econ. Rev. 83 (4), 792–810.

Cabrera, M., Fierro, J., 2013. Implicaciones ambientales y sociales del modelo extractivista en Colombia. En Garay L.J. et al. (2013) Minería en Colombia: Fundamentos para superar el modelo extractivista, Contraloría General de la República, Colombia.

Cadsby, C., Maynes, E., 1999. Voluntary contribution to threshold public goods with continuous provision: experimental evidence. J. Public Econ. 71, 53–73.

Castellanos, A., Chaparro-Narváez, P., Morales Plaza, C.D., Alzate, A., Padilla, J., Arévalo, M., Herrera, S., 2016. Malaria in gold-mining áreas in Colombia. Memorias Instituto Oswaldo Cruz 111 (1), 59–66.

Chaudhuri, A., 2010. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Exp. Econ. 14, 47–83.

Clements, T., John, A., Nielsen, K., An, D., Tan, S., Milner-Gulland, E.J., 2010. Payments for biodiversity conservation in the context of weak institutions: comparison of three programs from Cambodia. Ecol. Econ. 69 (6), 1283–1291.

Cochard, F., Willinger, M., Xepapadeas, A., 2005. Efficiency of Nonpoint Source Pollution Instruments: An Experimental Study. Environmental and Resource Economics 30, 393–422.

Cremers, L., de Theije, M., 2013. Small-scale gold mining in the Amazon. In: Cremers, L., Kole, J., de Theije, M. (Eds.), Small-Scale Gold Mining in the Amazon. The Cases of Bolivia, Brazil, Colombia, Peru and Suriname. Cuadernos del CEDLA No 26. Centre for Latin American Studies and Documentation, Amsterdam, The Netherlands.

Crompton, P., Ventura, A.V., de Souza, J.M., Santos, E., Strickland, G.T., Silbergeld, E., 2002. Assessment of mercury exposure and malaria in a Brazilian Amazon riverine community. Environ. Res. 90 (2), 69–75.

Croson, R, Marks, M.B., 2000. Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis. Experimental Economics 2, 239–259.

Dannenberg, A., Loschel, A., Paolachi, G., Reif, C., Tavoni, A., 2014. On the provision of public goods with probabilistic and ambiguous thresholds. Environ. Resour. Econ. 61 (3), 365–383.

Dawson, N.L., Segerson, K., 2008. Voluntary agreements with industries: participation incentives with industry-wide targets. Land Econ. 84 (1), 97–114.

Feigenberg, B., Field, E., Pande, R., 2013. The economic returns to social interaction: experimental evidence from microfinance. Rev. Econ. Stud. 80 (4), 1459–1483.

Feigenberg, B., Field, E., Pande, R., Rigol, N., Sarkar, S., 2014. Do group dynamics influence social capital gains among microfinance clients? Evidence from a randomized experiment in urban India. J. Policy Anal. Manag. 33 (4), 932–949.

Giordana, G., Willinger, M., 2013. Regulatory instruments for monitoring ambient pollution. In: List, J., Price, M. (Eds.), Handbook on Experimental Economics and the Environment. Edward Elgar Publishing, Cheltenham UK, Northampton MA, USA.

Giraldo, J., Muñoz, J.C., 2012. Informalidad e ilegalidad en la explotación del oro y la madera en Antioquia. Universidad EAFIT, Fundación Proantioquia.

Griffin, D., Husted, B., 2015. Social sanctions or social relations? Microfinance in Mexico. J. Bus. Res. 68, 2579–2587.

Hadi, A.B., Kamaluddin, A., 2015. Social Collateral, repayment rates, and the creation of capital among the clients of microfinance. Procedia Economics and Finance 31, 823–828.

Hayes, T., Murtinho, F., Wolff, H., 2015. An institutional analysis of Payment for Environmental Services on collectively managed lands in Ecuador. Ecol. Econ. 118, 81–89.

Hayes, T., Murtinho, F., Wolff, H., 2017. The impact of payments for environmental services on communal lands: an analysis of the factors driving household land-use behavior in Ecuador. World Dev. 93, 427–446.

Hentschel, T., Hruschka, F., Priester, M., 2003. Artisanal and Small-Scale Mining. Challenges and Opportunities. International Institute for Environment and Development (IIED) and World Business Council for Sustainable Development (WBCSD), London.

Hilson, G. (Ed.), 2003. The Socio-Economic Impacts of Artisanal and Small-Scale Mining in Developing Countries. A.A. Balkema Publishers, Lisse, The Netherlands.

Hinton, J., 2005. Communities and Small-Scale Mining: an Integrated Review for Development Planning. Report to the World Bank (available on the web).

Holmstrom, B., 1982. Moral hazard in teams. Bell J. Econ. 13, 324–340.

Kaczan, D., Pfaff, A., Rodriguez, L.A., Shapiro, L., 2017. Increasing the impact of collective incentives in payments for ecosystem services. J. Environ. Econ. Manag. 86, 48–67.

Karlan, D., 2007. Social connections and group banking. Econ. J. 117 (February), F52–F84.

Kerr, J.M., Vardhan, M., Jindal, R., 2014. Incentives, conditionality and collective action in payment for environmental services. Int. J. Commons 8 (2), 595–616.

Kotchen, M., Segerson, K., 2019. On the use of group performance and rights for environmental protection and resource management. Proc. Natl. Acad. Sci. 116 (12), 5285–5292.

Ledyard, J., 1995. Public goods: a survey of experimental research. In: Kagel, J., Roth, A. (Eds.), Handbook of Experimental Economics. Princeton University Press.

Levi, M., 1988. Of Rule and Revenue. University of California Press, Berkeley.

Miceli, T., Segerson, K., 2007. Punishing the innocent along with the guilty: the economics of individual versus group punishment. J. Leg. Stud. 36, 81–106.

Midler, E., Pascual, U., Drucker, A.G., Narloch, U., Soto, J.L., 2015. Unravelling the effects of payments for ecosystem services on motivations for collective action. Ecol. Econ. 120, 394–405.

Moros, L., Vélez, M.A., Corbera, E., 2007. Payments for ecosystem services and motivational crowding in Colombia's Amazon Piedmont. Ecol. Econ. 156, 468–488.

Narloch, U., Pascual, U., Drucker, A.G., 2012. Collective action dynamics under external rewards: experimental insights from andean farming communities. World Dev. 40 (10), 2096–2107.

Narloch, U., Drucker, A.G., Pascual, U., 2017. What role for cooperation in conservation tenders? Paying farmer groups in the High Andes. Land Use Policy 63, 659–671.

Ostrom, E., 1990. Governing the Commons. The Evolution of Institutions of Collective Action. Cambridge University Press.

Poe, G., Schulze, W.D., Segerson, K., Suter, J., Vossler, C., 2004. Exploring the performance of ambient-based policy instruments when nonpoint source polluters can cooperate. Am. J. Agric. Econ. 5, 1203–1210.

Reichhuber, A., Camacho, E., Requate, T., 2009. A framed field experiment on collective enforcement mechanisms with Ethiopian farmers. Environ. Dev. Econ. 14, 641–663.

Saldarriaga-Isaza, A., Villegas-Palacio, C., Arango, S., 2013. The public good dilemma of a non-renewable common resource: a look at the facts of artisanal gold mining. Resour. Policy 38, 224–232.

Saldarriaga-Isaza, A., Arango, S., Villegas-Palacio, C., 2015. Phasing out mercury through collective action in artisanal gold mining: evidence from a framed-field experiment. Ecol. Econ. 120, 406–415.

Salk, C., Lopez, M.C., Wong, G., 2017. Simple incentives and group dependence for successful payments for ecosystem services programs: evidence from an experimental game in rural Lao PDR. Conservation Letters 10 (4), 414–421.
Sarmiento, M., Giraldo, B.H., Ayala, H., Uran, A., Soto, A.C., Martinez, L., 2013. Characteristics and challenges of small-scale gold mining in Colombia. In: Cremers, L., Kole, J., de Theije, M. (Eds.), Small-Scale Gold Mining in the Amazon. The Cases of Bolivia, Brazil, Colombia, Peru and Suriname. Cuadernos del CEDLA No 26. Centre for Latin American Studies and Documentation, Amsterdam, The Netherlands.
Segerson, K., 1988. Uncertainty and the incentives for nonpoint pollution control. J. Environ. Econ. Manag. 15, 87–98.
Segerson, K., 1998. Voluntary vs. Mandatory Approaches to Nonpoint Pollution Control: Complements or Substitutes? Working Paper. Department of Economics, University of Connecticut Storrs.
Segerson, K., 1999. Flexible incentives: a unifying framework for policy analysis. In: Casey, F., Schmitz, A., Swinton, S., Zilberman, D. (Eds.), Flexible Incentives for the Adoption of Environmental Technologies in Agriculture. Kluwer, Norwell, MA, pp. 79–95.
Segerson, K., 2013. Voluntary approaches to environmental protection and resources management. Annual Review of Resource Economics 5, 161–180.
Segerson, K., Wu, J., 2006. Nonpoint pollution control: inducing first-best outcomes through the use of threats. J. Environ. Econ. Manag. 51, 165–184.
Shoko, D., 2002. Small-scale and alluvial gold panning within the Zambezi Basin: an ecological time bomb and tinder box for future conflicts among riparian states. In: Paper Presented at the 9th Conference of International Association for the Study of Common Property. Victoria Falls, Zimbabwe.
Siegel, S., 2013. Community without solidarity: mercury pollution from small-scale mining and Colombia's crisis of authority. Community Dev. J. 48 (3), 451–465.
Sommerville, M., Jones, J.P., Rahajaharison, M., Milner-Gulland, E.J., 2010. The role of fairness and benefit distribution in community-based Payment for Environmental Services interventions: a case study from Menabe, Madagascar. Ecol. Econ. 69 (6), 1262–1271.
Spraggon, J., 2002. Exogenous targeting instruments as a solution to group moral hazards. J. Public Econ. 84, 427–456.
Stiglitz, J., 1990. Peer monitoring and credit markets. World Bank Econ. Rev. 4 (3), 351–366.
Suleiman, R., Rapoport, A., 1992. Provision of step-level public goods with continuous contribution. J. Behav. Decis. Mak. 5, 133–153.
Suter, J.F., Vossler, C.A., Poe, G.L., Segerson, K., 2008. Experiments on damage-based ambient taxes for non-point source polluters. Am. J. Agric. Econ. 90 (1), 86–102.
Suter, J.F., Segerson, K., Vossler, C.A., Poe, G.L., 2010. Voluntary-threat approaches to reduce ambient water pollution. Am. J. Agric. Econ. 92 (4), 1195–1213.
Swenson, J., Carter, C., Domec, J.C., Delgado, C., 2011. Gold mining in the Peruvian Amazon: global prices, deforestation and mercury imports. PLoS One 6, 1–7.
Velez, M.A., 2011. Collective titling and the process of institution building: the new common property regime in the Colombian pacific. Hum. Ecol. 39, 117–129.
Vossler, C.A., Poe, G.L., Segerson, K., Schulze, W.D., 2006. Communication and incentive mechanisms based on group performance: an experimental study of nonpoint pollution control. Economic Inquire 44, 599–613.
Wade, L., 2013. Gold's dark side. Science 341 (6153), 1448–1449.
Xepapadeas, A., 1991. Environmental policy under imperfect information: incentives and moral hazard. J. Environ. Econ. Manag. 20, 113–123.