Land Use Policy 86 (2019) 218–228

Contents lists available at ScienceDirect

Land Use Policy

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

Effectiveness and equity of Payments for Ecosystem Services: Real-effort experiments with Vietnamese land users

Lasse Loft[a,*], Stefan Gehrig[b], Dung Ngoc Le[c], Jens Rommel[d]

[a] Leibniz Centre for Agricultural Landscape Research (ZALF), Eberswalder Str. 84, 15374 Müncheberg, Germany
[b] Independent Researcher, Berlin, Germany
[c] The Netherlands Development Organisation (SNV), Viet nam
[d] Swedish University of Agricultural Sciences (SLU), Sweden

ARTICLE INFO

Keywords:
Social justice
Fairness
Experimental economics
Policy instruments
Lab-in-the-field experiments
Forest conservation
Social norms
Coordination
South East Asia

ABSTRACT

Payments for Ecosystem Services (PES) are widespread in conservation policy. In PES, environmental effectiveness and social equity are often perceived as conflicting goals. Empirical studies on the relationship between popular design features, such as payment differentiation and payment conditionality, and effectiveness and equity are scarce. Further, they struggle with measuring and separating ecological and equity outcomes. In this study, we combine two incentivized lab-in-the-field experiments with 259 land users from eight villages in North-Western Vietnam to assess both individual conservation effort and community-level equity perceptions under four different PES designs. Effort is measured in a real-effort task with real-world environmental benefits; equity perceptions about payment designs in the real-effort task are measured in a coordination game. We demonstrate that payment design affects both effort and equity perceptions. Payments which are differentiated and are solely conditional on individuals' contributions of effort are perceived as most equitable. They are also more effective in motivating conservation effort than other designs, although the differences are small and not significant for all comparisons. By working out the positive correlation of effectiveness and equity across the four payment schemes, we show that these objectives are not necessarily conflicting goals in incentive-based conservation policy. Further, we can show that women exert greater conservation efforts. We discuss how greater equity and effectiveness could be achieved with reforms towards more input-based distribution criteria in Vietnam's PES legislation and the limitations and opportunities of the experimental paradigm for research on PES.

1. Introduction

Global efforts for the conservation of biodiversity and ecosystem services, such as carbon sequestration and storage or water filtration, call for a diverse set of policy interventions at various governance levels. Over the past years, incentive-based approaches such as Payments for Ecosystem Services (PES) have developed into a widespread conservation policy instrument (Muradian et al., 2010; Sattler and Matzdorf, 2013; Schomers and Matzdorf, 2013; Vatn, 2015; Wunder, 2015). Design and implementation of these instruments often encompass both, ecological effectiveness and social equity objectives (He and Sikor, 2015; Law et al., 2018; Loft et al., 2017a; Sills et al., 2017). In the literature on social equity in conservation initiatives, most studies report that land and natural resource users suffer from inequitable procedures or distributions (Friedman et al., 2018). Importantly, equity may be linked to environmental outcomes in different ways: On the one hand, achieving both goals simultaneously with a limited budget is often viewed as a trade-off (Chu et al., 2019; Martin et al., 2014a; Pascual et al., 2010; Wunder et al., 2018). Thus, given a fixed budget, equity and environmental effectiveness should be inversely correlated. On the other hand, it has been argued that equity itself affects effectiveness, with a positive causal effect of fair distributions to effective conservation (see Bennett, 2016; Law et al., 2018; Pascual et al., 2014). In this study, we focus on the former type of link. We follow the call that there is a "need for more research on the potentially important link between social equity and environmental effectiveness" (Engel, 2016, p. 138) and assess different designs of incentive-based conservation policies separately in both dimensions, their effectiveness to induce conservation effort and their perceived equity.

The concept of equity (fairness, justice) implies much more than equality. For example, it incorporates whether people are seen as accountable for the conditions they experience or the opportunities they

* Corresponding author.
E-mail addresses: lasse.loft@zalf.de (L. Loft), stefan-gehrig@t-online.de (S. Gehrig), dlengoc@snv.org (D.N. Le), jens.rommel@slu.se (J. Rommel).

https://doi.org/10.1016/j.landusepol.2019.05.010
Received 15 February 2019; Received in revised form 5 May 2019; Accepted 5 May 2019

have (e.g., instead of having luck or profiting from others' efforts; Konow, 2003). Equity in PES is often defined along four dimensions: *context* as the acknowledgment of local conditions; *procedure* as the participation in implementation; *recognition* as the respect for all stakeholders' knowledge and values; and *distribution* in terms of costs and benefits occurring to different stakeholders (Brown and Corbera, 2003; Corbera et al., 2007; Friedman et al., 2018; Martin et al., 2014a, 2014b; McDermott et al., 2013; Pascual et al., 2010, 2014). In particular, the latter has been the focus of much debate, because the perception of what constitutes a fair distribution of the costs and benefits of conservation among local program participants differs within and between communities and stakeholder groups (e.g., Hayes and Murtinho, 2018; He and Sikor, 2015; Loft et al., 2017b; Narloch et al., 2013). This can lead to contradictory conclusions about which benefit distribution is fair. Indeed, Myers et al. (2018) identify benefit sharing and its alignment with local notions of justice as a primary source of conflict in PES policies. To navigate the potential trade-off between benefit sharing and environmental effectiveness in PES, various design features of PES schemes have been discussed. Fundamental design questions are whether and how payments should be differentiated between program participants and to which extent they should be conditional on performance (Engel, 2016; Luttrell et al., 2013; McDermott et al., 2013; Pascual et al., 2010; Wunder et al., 2018). PES programs may pay fixed rates to local ecosystem service providers or adjust payments by different factors. Payments could, for example, be adjusted relative to ES providers' costs (e.g., their opportunity costs of refraining from resource extraction, i.e., "compensation") or the actual provision of ES (i.e., the degree of ecological improvement caused by the intervention). Importantly, the details of these design features can affect ecological and social outcomes of PES (Ezzine-de-Blas et al., 2016; Wunder et al., 2018). While it is broadly acknowledged that differentiation and conditionality influence social and ecological outcomes of PES policies, there have been few attempts to disentangle the equity and effectiveness consequences of contract design features (Calvet-Mir et al., 2015; Halpern et al., 2013; Pascual et al., 2010). This is probably partially due to the challenge of cleanly measuring and causally attributing equity and effectiveness outcomes in case studies and observational research.

In this study, accordingly, we explore environmental effectiveness (research question 1) and social equity perceptions (research question 2) of different payment contracts separately by using economic experiments with a population of PES recipients in North-Western Vietnam. In particular, we investigate undifferentiated egalitarian payments, payments conditional on individual costs and efforts, and payments conditional on actual provision. We also investigate payments that differentiate between recipients, but that do so randomly and without being conditional on salient outcome measures. By subsequently aggregating the separate results on social equity and environmental effectiveness from two related experiments, we will thus also have an empirical basis to comment on the potential trade-off between both outcome dimensions (research question 3), as proposed in the literature (Chu et al., 2019; Law et al., 2018; Narloch et al., 2013; Wunder et al., 2018). We introduce a novel combination of economic lab-in-the-field experiments. In line with the calls for more interdisciplinary research on justice in conservation policies (Friedman et al., 2018) and more evidence-based analyses of equity in PES (Calvet-Mir et al., 2015; Martin et al., 2014a; Pascual et al., 2010, 2014; Vatn, 2010), we implement paradigms from experimental and behavioral economics with forest resource users in the field. This gives us the opportunity to compare different PES benefit sharing schemes in the same population, instead of comparing case study reports which differ over many contextual factors. We can draw on a large research tradition in economics that links fairness considerations or social norms to economic behavior (e.g., Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Krupka and Weber, 2013).

Our study, thus, augments the growing literature on lab-in-the-field experiments on the effectiveness of PES (Andersson et al., 2018; Gatiso et al., 2018; Handberg and Angelsen, 2019; Narloch et al., 2012; Reutemann et al., 2016; Salk et al., 2017) and builds on the premise that experimental economics is well-suited to analyse effects of institutions and policies on pro-environmental behavior (Rommel, 2015).

In the next section, we introduce the study site and population as well as the current PES regime in the area. Then, we turn to our experimental methods. After reporting experimental results, we close with a discussion.

2. Local background

In this section, the national and regional background on payments for forest conservation in Vietnam are explained in more detail (2.1) with further elaborations on the current benefit sharing rules of the scheme (2.2).

2.1. National and regional context

A variety of PES definitions and typologies exist. A core element of most PES definitions is the provision of positive economic incentives conditional upon the supply of well-defined ecosystem services (or activities thought to yield well-defined ecosystem services) (Muradian et al., 2010; Schomers and Matzdorf, 2013; Wunder, 2015, 2005). In Vietnam, the national Payments for Forest Ecosystem Services scheme (PFES) was piloted between 2008 and 2010 and has been implemented nationwide since 2011 (McElwee et al., 2014). PFES targets four broad categories of ecosystem services provided to different non-governmental ecosystem services buyers: (1) watershed protection; (2) protection of natural landscape beauty and conservation of forest biodiversity; (3) forest carbon sequestration and retention; (4) provision of spawning grounds and use of water from forests for aquaculture. In contrast to some definitions of PES (e.g. Wunder, 2005, 2015), PFES participation of ecosystem service providers and users is mandatory by law. Currently, hydropower plants and water supply companies are the main buyers of ecosystem services. Their payments are delivered to government agencies and then disbursed to ecosystem services providers, i.e., land users. Contrary to some other PES schemes, land users are not paid for the outcome in terms of provided ecosystem services. Instead, they are financially compensated for undertaking active forest protection, such as firefighting and organizing "forest protection groups". Among the main beneficiaries of PFES payments are households and communities with rights to forest land. Currently, on the national level, 355,000 households and community groups, managing 3.5 million hectares of forest (25% of total forest area in the country), receive PFES payments (Vietnam Forest Protection and Development Fund (VNFF), 2015). In our study province of Dien Bien in North-Western Vietnam (Fig. 1), large parts of the forests are managed by individuals, households, and communities. In Dien Bien, 12 million USD of PFES revenues had been collected from buyers since 2011, of which 5 million USD had been distributed to villagers as providers of ecosystem services at the time of data collection.

We selected eight villages (Fig. 1) of three different ethnicities (Thai, Hmong, Khang) that were part of an earlier study which analysed the consideration of equity in the design and implementation of PFES in Vietnam (Loft et al., 2017b; Supplementary material, Table S1). This allowed us to conduct the experimental study with a solid understanding of the local context. Originally, the villages had been selected by the researchers based on representativeness of the social, political, and economic characteristics of the region (e.g., ethnic diversity, poverty). The villages are also well-suited for our research questions on design features of payment differentiation and conditionality: Loft et al. (2017b) showed that, in these villages, perceptions of distributional equity in PFES are mainly oriented along the concepts of egalitarianism (equal payments) and compensation (merit-based payments), while alternative design features like needs-based (pro-poor payments) or status-based approaches play only a small role.

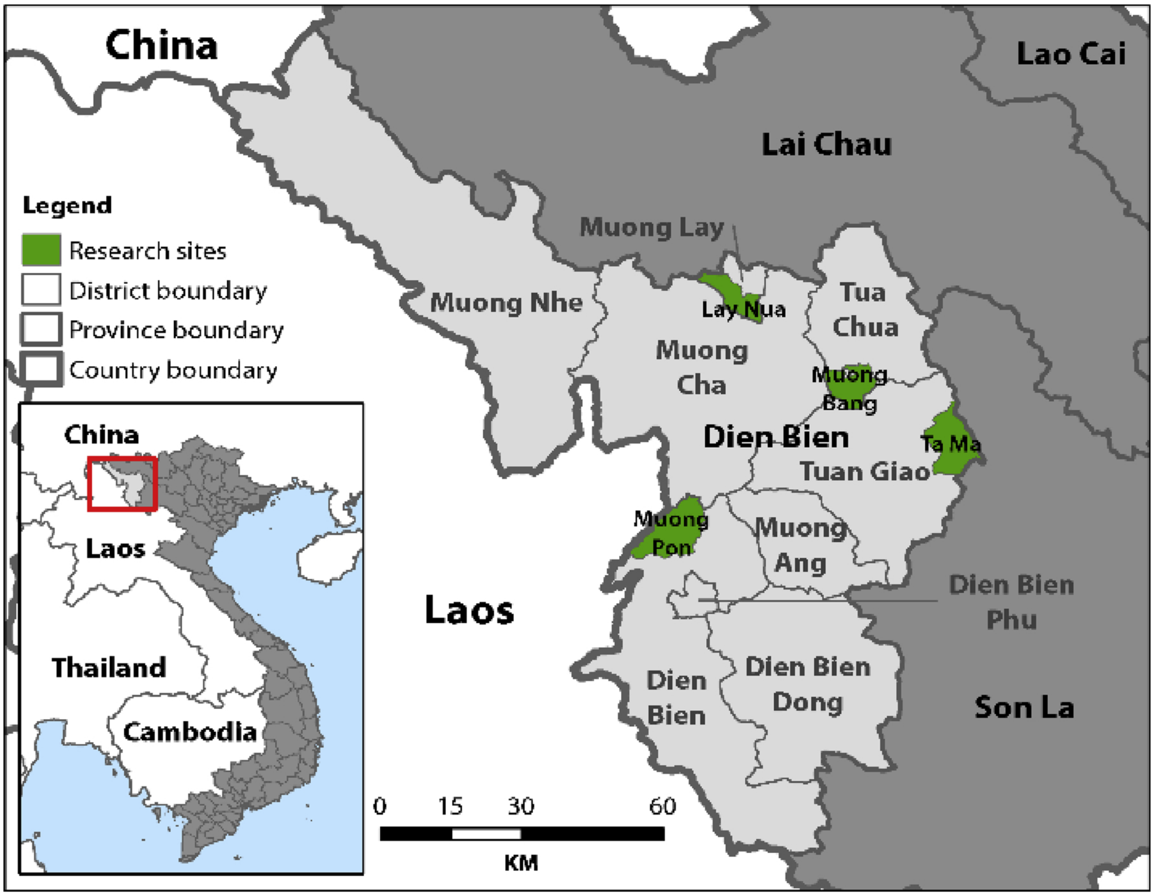

Fig. 1. Study area. Within each research site (green), two villages took part in the study (for interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

2.2. Current payment scheme

Under PFES, the size of payments to land use rights holders is determined by a sequence of formulae, laid out in Vietnam's national legislation. The net revenues of payments collected from watershed service buyers are disbursed to ecosystem services providers. So-called "K-coefficients" are being applied to differentiate payments based on forest quality, classification, and accessibility. K1 differentiates forest condition (rich, medium, or poor); K2: forest function (special-use, protection, or production forest); K3: origin of the forest (natural or plantation); and K4: level of difficulty of forest protection (considering social and geographic factors). The purpose of these K-coefficients is to reward "who is doing the best job in providing the ecosystem services" and to ensure that "those who keep the forest well, deserve high payments" (personal communication, VNFF). K1–3 are indicators for the quality of the forest, used as a proxy for the provision of ecosystem services. Theoretically, the application of K1–3 allows for a differentiation of payments based on the result of forest protection. Thus, they reflect a distribution that rewards the output or "actual provision" of ecosystem services. The application of K4 allows for a differentiation of payments according to the difficulty of forest protection based on social and geographic factors. It thereby takes into account the compensation of work effort of ecosystem services providers. For example, if the forest areas to be protected are distant from a settlement area, the level of payment should be adjusted to compensate for time and transportation costs. Thus, K4 is a merit-based distribution that compensates those who put the most work into forest protection. However, the application of these K-coefficients is currently often suspended in practice due to their vague definition, administrative capacity constraints, high transaction costs, and a lack of guidelines and indicators for implementation. Due to the practical simplicity in implementation, equal payments are often being distributed (Loft et al., 2017b; Pham et al., 2013). Our results on social and ecological consequences of different payment design features in a PFES-targeted population hence will also speak to the discourse of payment design and implementation in the area.

3. Methods

Our study consisted of two lab-in-the-field experiments, both of which were conducted in each of the eight study villages (see Supplementary material, Table S1 for characteristics of villages and sessions). Note that all data and code are available for a full reproduction of all results and figures presented in this paper in the Code Ocean capsule https://codeocean.com/capsule/1108125.

The first experiment (Experiment 1) was a real-effort task on the impact of different payment schemes on actual conservation effort, i.e., environmental effectiveness (Table 1). Conservation effort was measured as performance in the real-effort task in which real public goods for forest conservation were produced. In addition, we conducted a coordination game to elicit "equity scores" as a community-level measure of perceived fairness for each of the four payment schemes (Experiment 2). Each experiment used a different group of subjects from the same village, i.e., subjects could participate in only one of the experiments, and we ensured separation to prevent communication. We analyze effectiveness in Experiment 1 (Research question 1) and equity in Experiment 2 (Research question 2) separately, before looking at their association (Research question 3; see Fig. 2).

Although we informed village heads about our study and received their consent, participants were unaware of the details of the planned

Table 1
Treatments in the real-effort task.

Treatment	Reference in the PES literature	Payoff function	Sample size
Merit input (differentiated, conditional payment)	Distribution should be proportional or relative to the contribution of the stakeholders. The distribution compensates costs and input such as work effort (Pascual et al., 2010; Sattler and Matzdorf, 2013)	$P_i = x_i * A$ (piece rate)	43
Merit output (differentiated, conditional payment)	Distribution rewards output, such as monitored change in the provision of ES ("actual provision") (Pascual et al., 2010; Sattler and Matzdorf, 2013).	$P_i = x_i * A + 5 * A$ (piece rate)	44
Fixed equal (undifferentiated, unconditional payment)	Distribution is equal among all providers of a service independent of the cost and level of service provision (Pascual et al., 2010).	$P_i = B$ (flat rate)	42
Fixed individual (differentiated, unconditional payment)	Distribution is differentiated, but without salient justifiable criteria. For example, payment distribution via auctions might get perceived as random allocation by people who lack understanding of the auction mechanism (Leimona and Carrasco, 2017).	$P_i = C_i$ (flat rate)	44

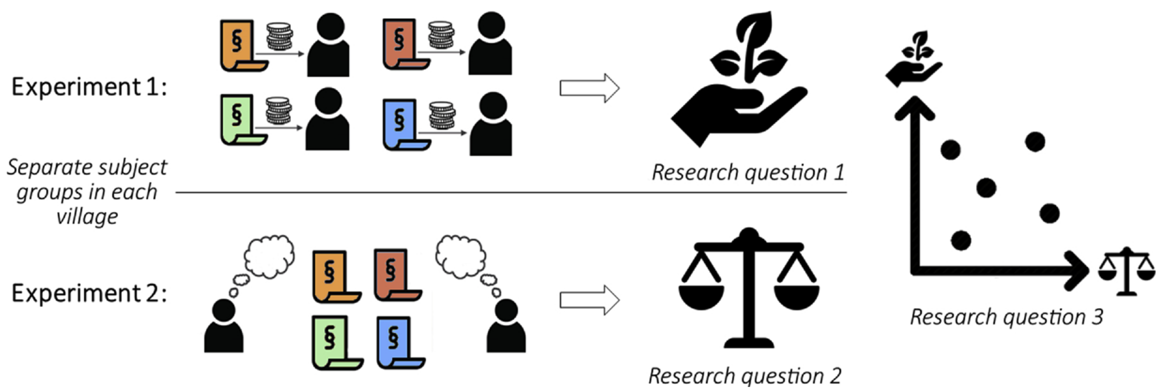

Fig. 2. Empirical strategy and research questions. Differently colored contracts represent different payment schemes that vary in their design features with respect to payment differentiation and conditionality. Experiment 1 focused on ecosystem service provision conditional on contract design (real-effort task). Experiment 2 focused on beliefs about the social appropriateness of the different contract designs (coordination game).

experiments to prevent strategic communication prior to implementation. To limit cross-communication among participants, for each experiment we conducted only one session in a given village. We ensured that participants of the two experiments did not communicate verbally during a session.[1] In total, 259 subjects participated in the experiments (176 in Experiment 1[2], and 83 in Experiment 2), of which 139 (54%) were female. Sampling of participants within villages was initiated via a public announcement, with subsequent randomization of people who showed interest into participation (yes or no) and experiment (1 or 2). Note that due to our rather small sample sizes, our design allows us to detect only medium-sized and large effects at conventional levels of alpha = 5% and beta = 20%.

Instructions were given in Vietnamese, and, if applicable, in the local language. Subjects were free to leave at any time, all decisions were made in private, payment distribution followed a double-blind process, and there was no deception in the experiments. The show-up fee was 2.19 USD, which approximates a daily wage in the region. On top of that, subjects earned on average 2.11 USD (SD = 0.73 USD) in Experiment 1 (90 min) and 1.00 USD (SD = 0.39 USD) in Experiment 2 (45 min).[3] After completion of the task and before payments, subjects took part in a brief survey on socio-demographics, conservation behavior and experiment-related controls. We gathered all participants for a debriefing after the experiments, in which we explained the study purpose. There was time to discuss the experiment with the researchers and to ask questions about the study.

3.1. Experiment 1: the real-effort task

In this section, the design of the real-effort task (3.1.1), its different payment treatments (3.1.2), and the experimental procedure in the field (3.1.3) are explained.

3.1.1. Task design

As common in experimental economic research on the role of incentives (Charness et al., 2018), we used a real-effort experiment to determine environmental effectiveness under different payment schemes. Environmental effectiveness was measured as conservation effort in the task. The task required physical effort and produced real-world environmental benefits, which distinguishes it from abstract pen-and-paper experiments on PES in which subjects invest in or harvest hypothetical resources and where their payoffs are exclusively non-environmental (e.g., Andersson et al., 2018; Handberg and Angelsen, 2019; Narloch et al., 2012). Real-effort tasks can add validity to experimental research concerned with behavior in a particular applied context (Charness et al., 2018), and they have recently become popular in conservation research with natural resource users in the field (Jindal et al., 2017; Nelson et al., 2018).

Lab-in-the-field experiments that produce real-world public goods are still only scarcely employed (but see Gatiso et al., 2018 for experiments that fund schools), in particular in combination with real-effort tasks (Kerr et al., 2012). Consequently, we provide the first experimental study on PES which encompasses real effort of real resource users and real environmental benefits.

In the present task, subjects had 45 min to fill small degradable

[1] In both experiments, subjects were encouraged not to talk and indeed did not talk, aside from occasional laughter or moaning (as reported by local assistants and the Vietnamese co-author who was part of the field research team). In Experiment 2, subjects were seated in a way that they could not observe others' decisions.

[2] We excluded three subjects of Experiment 1 from the analysis, as we could not rule out that they had listened to the instructions or talked to participants of the other experiment. This resulted in 173 subjects for our analysis of Experiment 1 and 256 in total.

[3] All payments were made in Vietnamese Dong. For better readability, we report all monetary values in US Dollar. At the time of the study 10,000 Dong were equal to 0.44 US Dollar (1 USD = 22,800 Dong).

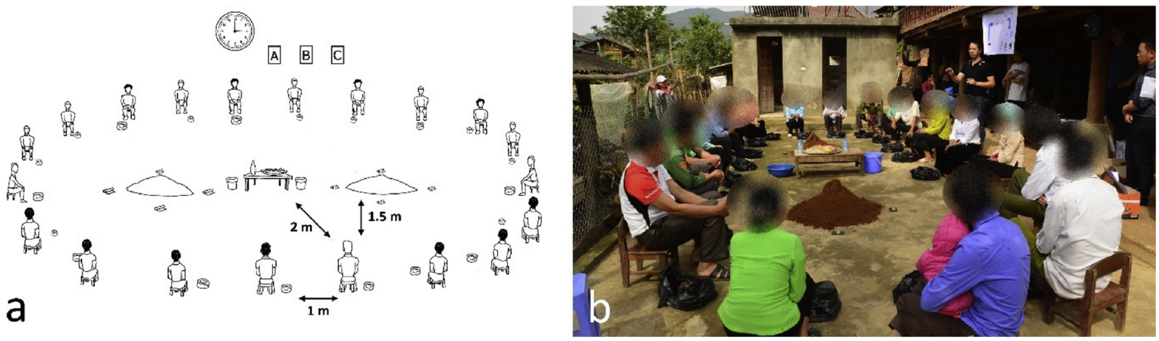

Fig. 3. Real-effort task set-up for the production of soil bags (a. concept, b. example of actual implementation). Subjects were seated in a circle in around the resources they needed for work (piles of bare soil, surrounded by empty soil bags) and their leisure option (snacks, drinks, buckets for hand washing). Bags had to be filled with soil. We asked subjects to store all produced bags in a larger individual bag next to each seat. Subjects were free to stop working and move around (except for leaving the experimental setup), and to consume snacks at any time. They were not allowed to talk. A large watch indicated the remaining time. Sheets, titled A, B and C, were visible to all participants, and their backsides contained the randomly implemented payment rate for each treatment (see 3.1.2), which were revealed after the task.

plastic bags (3 x 9 cm) with nutritious soil (see Fig. 3 for setup). The performance in the task was easy to measure (number of bags produced) and could be observed under different payment schemes, as reflected in our treatments (Table 1). Note that such soil bags are commonly used by conservation NGOs in the region. Just as in our task, organizations pay people for filling these soil bags which are later used for afforestation (Supplementary material). We informed subjects prior to the task that all bags they produce would be used for afforestation in the region; they were collected by forest officials immediately after the task. Hence, subjects knew that they were producing a real-world environmental public good and real environmental concerns should therefore affect behavior in the task, besides any payments.[4] This adds an important layer of realism to the experiments, which is absent in lab-in-the-field experiments purely based on social dilemma games. In such games, researchers measure cooperation which does not necessarily equal pro-environmental behavior (Gehrig et al., 2019; Torres-Guevara and Schlüter, 2016). Contrary to these games, the social dilemma does not arise in small groups within the game, but on the level of the real-world conservation public good (here, forests). The task had additional desirable features that would reflect real-world conservation activities: it was tedious and repetitive, but it did not require any special skills that could vary strongly among subjects (we also asked and controlled for previous experience in a survey question). As part of PFES policies, people in our study population must also actively engage in effort for forest protection, amongst others through monitoring and reporting illegal logging activities and firefighting (Loft et al., 2017b; Pham et al., 2013). Also in other instances, for example in the Sloping Land Conversion Program in China (He and Sikor, 2015), active afforestation is a fundamental part of PES schemes. Thus, we argue that an experiment which measures active effort as our real-effort task (instead of just the costs of foregone benefits from resource extraction) is suited for research on PES generally and for the PFES case in particular.

Subjects in real-effort tasks may exert effort because of boredom or a lack of attractive alternative things to do, i.e., often, there are no opportunity costs in an experimental setting and this can affect performance (see Eckartz, 2014; Erkal et al., 2018). For example, curiosity for the task, rather than experimentally controlled incentives (money, conservation) could affect performance. To reduce such confounding effects, we offered sweets, snacks, and water to subjects on a table as an alternative to working (Fig. 3). We removed snacks immediately after the task, to prevent that subjects would plan to eat and drink afterwards. This was common knowledge before the start. Snacks had a value of approximately 5 USD per village. This outside option was extensively used in all villages but not by all subjects. Several subjects took breaks for several minutes. Although we have no data on the time used for consumption per session or individual, we recorded that the latest initial "snack break" taken in any of the village was after 25 min, and the median initial break was after 15 min. This value does not correlate with mean village-level performance (Supplementary material).

3.1.2. Treatments

We randomly assigned participants to four different treatments, i.e. payment schemes in each village. We defined the treatments based on commonly discussed PES design features of payment differentiation and conditionality (Table 1). In the *merit input* treatment, subjects were paid solely based on their performance. In the *merit output* treatment, the individual containers already contained five soil bags to mimic heterogeneity in initial natural resource stocks that cannot be influenced by forest users themselves, e.g., better initial forest quality or the frequency of poaching by outsiders.[5] In the *fixed equal* payment, all participants received the same fixed payment, independent of performance. This treatment resembles unconditional egalitarian payments, i.e., everyone received the same money. In contrast, in the *fixed individual* treatment, there was variation in fixed payments on the individual level.

In Table 1, P_i is the payoff of subject i, and x_i the numbers of bags produced (i.e., performance) of subject i. A, B, and C_i are payment rates, which were randomly drawn between villages. We were interested in differences between the type of incentive in the payment schemes, not the impact of the magnitude of payment rates. We thus randomized the payment rates:

- A as piece payment rate for the *merit input* and *merit output* treatments,
- B as flat payment rate for the *fixed equal* treatment, and
- C_i as individual flat payment rate for each subject i in the *fixed individual* treatment

Payment rates were generated by determining one of the two

[4] This is mirrored in lab research in behavioral economics. If subjects perceive the tasks as meaningful, they exert greater efforts, independent of monetary rewards (e.g., Ariely et al., 2008; DellaVigna and Pope, 2018; Kosfeld et al., 2017). This is relevant for our study, because environmental conservation efforts in particular are often driven by the pursuit of meaning and moral satisfaction (see Rode et al., 2015).

[5] We chose this design for the *merit output* treatment, because it is easy to understand and captures the basic idea of external influences on performance. Alternatively, we could for example have introduced a probabilistic deduction or addition of bags.

possible payment rates with a coin flip before the task started but remained hidden behind visible sheets until after the task (Fig. 3) and should, thus, not affect effort differently between villages[6] . Subjects knew about the two possible outcomes to rule out that they formed beliefs about extremely high or low wage rates, which could have led to heterogeneity in beliefs and, subsequently, effort. To control for heterogeneity in expectations, we also elicited subjects' beliefs about their payment rates before the start of the task.

3.1.3. Procedure

In each village, between 20 and 28 subjects took part in Experiment 1 (Supplementary material, Table S1). Subjects were randomly assigned to one of the four treatments (Table 1). After a general welcome and introduction to the task, subjects received treatment-specific written instructions. Payments were individually and privately disclosed and explained in written form and with visual and verbal aids (see Supplementary material for instructions). Hence, subjects in Experiment 1 were not aware of other treatments than their own, which was meant to reduce experimenter demand effects (Zizzo, 2010). To nevertheless attain that the egalitarian and the differentiated flat rate payments (*fixed equal* and *fixed individual*, respectively) were perceived differently by subjects we framed all four treatments as group-based: we told subjects that they were in a group with some other people from their community who would receive the same instructions and type of payment as they do. Subjects were explicitly told that it was completely up to them how much they would work in the 45 min which would be available. They were also informed that the number of produced bags could not be connected to their names and identity. During the experiment, subjects could observe each other, but they were not allowed to talk. Although they could freely move, e.g., to wash their hands, relax, or consume snacks, subjects were not allowed to leave the experimental setting completely. A large watch was visible to all subjects and indicated the time. We ensured that subjects did not feel observed while working, by having only two of the six assistants present during the experiment (to prevent cheating and talking). The foreign researchers and fellow villagers were not present during the task. After 45 min, individual containers were collected, bags were counted, and subjects took part in a small survey (Supplementary material). Table 2 shows summary statistics for the participants in Experiment 1.

3.2. Experiment 2: the coordination game

In Experiment 2, our goal was to assign an "equity score" to each treatment and village that would reflect the perceived fairness of the payment scheme in a community at the time of experimentation. Experiment 2 was a coordination game. Coordination games are games in which the payoff of an individual depends on the frequency with which others choose the same option, but where, once coordinated, individuals have no reason to change their strategy. Instead of imposed definitions of equity, we experimentally elicited fairness perceptions about the four incentive mechanisms (Table 1) in the same community where the real-effort task was conducted, but with different subjects. The rationale was that distributional fairness norms may exhibit local variation, even on the relatively small scale of our study, which included three different ethnicities. Thus, we chose a "norm elicitation" method that allowed equity scores to vary between villages but that defined them on the village level, rather than the individual level. We define equity perceptions as a normative judgment and thus a property of the community that shares social bonds rather than a property of an individual. To measure norms of equity, we adapted an approach from the experimental economics literature: Krupka and Weber (2013, p. 499) define social norms as "collective perceptions, among members of a population, regarding the appropriateness of different behaviours." According to their definition, social norms create focal points to enhance coordination and cooperation in social interactions. Against this background, social norms are a "correlation device" to align mutual expectations about others' behavior in a community (see also Gintis, 2009). A community-level equity norm should enable people to predict others' equity perceptions about certain actions and behaviors. Thus, as opposed to survey measures which elicit data in isolation, a coordination game should be an appropriate approach to measure community-level norms, which, in our case, concerned the fairness of different PES schemes.

Following Krupka and Weber (2013), subjects had to rate the social appropriateness of each of the four treatments, i.e. payment schemes on a four-point scale from "very socially inappropriate" to "very socially appropriate". They were asked to picture a scenario in which individuals would face the real-effort task from Experiment 1 and would be paid by an organization according to each of the four different payment schemes. We described appropriateness further as being "consistent with moral or fair behavior" and as being the "right" thing to do.

Subjects received a payoff of 0.44 USD (10,000 Dong) for each response that would match the modal rating, i.e., the rating that most others also chose, among all other participants of Experiment 2 in their village. Hence, subjects had an incentive to state their true beliefs on what they think the plurality (relative majority) perceives as appropriate or fair and hence on what constitutes the shared social equity norm. Note again that this method does not view social norms as an aggregate of individuals' characteristics, but as a collective community characteristic that emerges through mutual expectations. Each subject rated each treatment (Table 1), i.e., everybody had to give four responses. Based on this data, we calculated an equity score for each of the four payment schemes and each village by a weighted summation of responses (Supplementary material), closely following the procedure of Krupka and Weber (2013): the equity score was defined as the sum of all "very inappropriate," "somewhat inappropriate," "somewhat inappropriate," and "very appropriate" ratings multiplied by -1, $-1/3$, $1/3$, and 1, respectively. Thus, an equity score of zero indicates "neutral" appropriateness.

In each village, either ten or eleven subjects took part in Experiment 2. We explained to them the coordination game and the real-effort task, because their choices were about the appropriateness of the type of payment that were provided in the latter. All subjects received written instructions, verbal and visual aids as a group, as well as personal support if questions remained. After a demonstration and practice round with an example, subjects rated the four treatments in private (see Supplementary material for instructions and materials).

4. Results

In this section, we first separately analyze performance in the real-effort task (4.1) and equity perceptions elicited in the coordination game (4.2), before combining the results in a joint analysis (4.3).

4.1. Experiment 1: effectiveness

Subjects invested considerable effort into production of soil bags for afforestation, as indicated by an average of 34.8 produced bags (SD = 13.8, median = 33) in the 45 min to perform the task. Effort aggregated over villages was approximately normally distributed (Supplementary material, Fig. S1). Only two subjects did not produce any bags within the 45 min (both in treatments with unconditional payments). The maximum number of bags produced was 80 (Fig. 4).

[6] The two possible payment rates for piece rate treatments (A) were 1,000 and 1,500 Dong per bag (0.04 and 0.07 USD, respectively). The two possible payment rates for the flat rate treatments (B and C_i) were 40,000 and 60,000 Dong (1.75 and 2.63 USD, respectively). Note that the initial endowment with five bags in *merit output* can be viewed as an increase in the show-up fee, i.e., an "unearned" and "undeserved" additional endowment.

Table 2
Summary statistics of subjects' characteristics in Experiment 1.

Variable	Description	n	Mean	SD	Min	Max
Forest group	Member of a forest protection group (1 = yes; 0 = no)	173	0.34	–	0	1
Male	Gender is male (1 = yes, 0 = no)	173	0.46	–	0	1
Income (USD)	Average daily income in USD	172	1	0.78	0.09	4.39
Age	Age in years	172	36.35	13.54	17	73
Years of education	Years of formal education (e.g., nursery, schooling)	173	5.07	4.31	0	15
Children in household	Number of children (15 years or younger) living in the household	172	1.66	1.32	0	6
Wealth index	First principal component of principal component analysis on household assets (motorcycle, bicycle, gas stove, cell phone, TV, fridge; proportion of variance explained = 0.345)	173	0	1	−0.98	2.6
Experience	Has prepared soil bags before (1 = yes; 0 = no)	170	0.23	–	0	1
Liked snacks	"I like the food and drinks that were offered." (1 = agree; 0 = neutral; −1 = disagree)	172	0.13	0.85	−1	1
Rate expectation	Indicates whether the participant has stated an expected wage rate above (= 1) or below (= 0) the expected value of the payment rate	172	0.65	–	0	1

Notes: Sample sizes vary slightly due to survey non-response. Details on the construction of the wealth index are provided in the supplementary material.

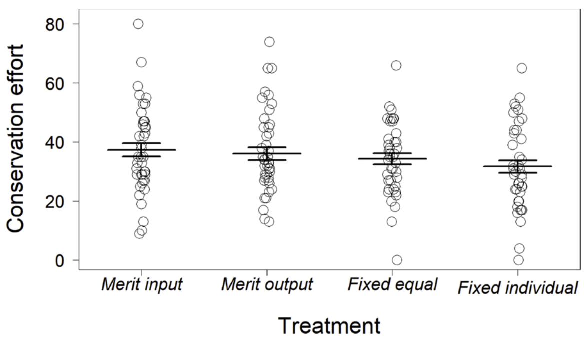

Fig. 4. Conservation effort (i.e., soil bags produced in the real-effort task) in Experiment 1, comparison over treatments (i.e., payment designs; for explanations, see Table 1). Each point is one observation (one subject). Horizontal lines indicate means and standard errors.

There was large between-village heterogeneity in performance (Supplementary material, Figs. S2, S3), with the mean village performance ranging from 21.8 to 50.9. As village variation is not our primary interest, we control for village in all subsequent statistical analyses (see Supplementary material for a short discussion on potential drivers of between-village heterogeneity). When including village as a predictor, an overall effect of treatment on performance in the task is indicated (ANOVA, $F_{3,\ 162} = 2.27$, $p = 0.082$; Fig. 4).

To separate the effect of different treatments, we use treatment dummy variables in a fixed-effect model with conservation effort (the number of bags produced) as a dependent variable (Table 3: Model 1; reference category: *fixed individual*). There are statistically significant treatment effects for the two merit-based payment schemes (*merit input* and *merit output*), when compared to the reference category. The effect of the *merit input* treatment is larger than the effect of the *merit output* treatment, but the two coefficients are not statistically different from each other. Neither are both different from the coefficient of *fixed equal* ($p = 0.14$ and $p = 0.26$, respectively, for estimates from Model 2 with switched reference category, see Code Ocean capsule).

In Model 2, we add socio-demographics and task-related survey variables to the models, as well as one indicator of real-life conservation behaviour (forest group). There are no issues with multicollinearity among the added independent variables (all pairwise correlations with Pearson $r < 0.35$). The estimated treatment effects do not substantially change in magnitude, and they maintain their level of statistical significance. We can conclude that conditional payments (both, based on compensation only, and based on actual provision of ES) lead to higher environmental effectiveness than unconditional, differentiated payments, while conservation effort under unconditional, undifferentiated payments (egalitarian payments) lies in-between and is not significantly different from any of the other treatments.

Turning to effects of individual-level covariates, there is a large gender effect on performance. Women, on average, produced five bags more than men. There are no large or statistically significant effects of other socio-demographics. The same applies to previous experience with the task, self-reported attractiveness of the snacks, i.e., an estimate of the magnitude of the individual opportunity costs of working, and expectations on the payment rate. Further, the proxy for real-world conservation behaviour, i.e. participation in a forest protection group, is not associated with conservation effort in the task.

Table 3
Models on conservation effort (i.e., soil bags produced in the real-effort task) in the real-effort task (Experiment 1) under different treatments (i.e., payment designs; for explanations, see Table 1). *Fixed individual* is the reference category. Models include village fixed effects. Standard errors are in parentheses.

	Model 1	Model 2
Merit input	5.66 (2.30)**	5.25 (2.43)**
Merit output	4.32 (2.29)*	4.38 (2.37)*
Fixed equal	2.53 (2.32)	1.62 (2.42)
Forest group		0.14 (2.01)
Male		−5.11 (1.87)***
Income (USD)		−0.80 (1.27)
Age		0.08 (0.08)
Years of education		0.06 (0.25)
Children in household		−0.39 (0.76)
Wealth index		0.22 (1.22)
Experience		0.63 (2.10)
Liked snacks		0.31 (1.04)
Rate expectation		−2.33 (1.93)
Constant	32.39 (3.56)***	34.32 (5.55)***
Village fixed effects	Yes	Yes
Observations	173	166
Log Likelihood	−658	−617
AIC	1,328	1,266

Notes: *p < 0.1, **p < 0.05, ***p < 0.01.

As robustness checks, we fit Model 2 also (1) including the three subjects we removed because they did not comply with the experimental protocol and (2) excluding outliers, defined as performance values that were exceeding 1.5 times the interquartile range of the aggregated outcome data. In both robustness checks, coefficients and standard errors remained virtually unchanged and p-values remained in the same intervals (see Code Ocean capsule).

4.2. Experiment 2: equity

In 65.4% of the cases, subjects matched (one of) their village's modal response(s) in Experiment 2, i.e., their response was in line with the plurality. In 81% of the cases, there was only one relative majority/modal response (i.e., there were no ties). In the other cases, all subjects who matched with one of the modal responses were paid. Remember that modal/plurality responses indicate a shared perception of what constitutes a "socially appropriate" and "fair" payment scheme. The results demonstrate considerable shared fairness perceptions among people from the same community.[7] Differences in ratings of payment schemes between communities were small in magnitude (Supplementary material, Fig. S4) and statistically not significantly different from each other (Kruskal-Wallis tests, all $p > 0.4$).[8] Thus, given the rather heterogeneous set of villages (e.g., regarding ethnicity or income, see Supplementary material, Table S1), it is fair to say that these norms are also shared across villages, at least at the regional scale of our study.

Overall, the *merit input* treatment was perceived as the most equitable (highest equity scores) across villages by far (median = 0.90, mean = 0.85), whereas the *fixed individual* treatment was ranked as the least equitable (median = −0.02, mean = −0.07), with the others ranked in between (Fig. 5). In six of the eight villages, the ranking of equity scores was consistent with the ranking of the pooled data (Supplementary material, Fig. S4). Equity scores differed significantly by payment scheme (Kruskal-Wallis test, $\chi^2 = 21.3$, df = 3, $p < 0.001$). All paired comparisons reveal differences between equity scores across payment schemes (Wilcoxon rank sum tests, all $p < 0.05$), except for the comparison of *merit input* and *fixed equal* ($p = 0.49$), as also apparent from Fig. 5. This implies that even *within* the two merit-based (conditional) payment schemes and within the two fixed (unconditional) payment schemes, there are large differences in what is perceived as more “socially appropriate” and “consistent with moral or fair behaviour.” Now, we turn to our primary aim, which was to relate equity perceptions about particular policies on the community level (Experiment 2) with individuals’ conservation behaviour under the respective policy (Experiment 1).

4.3. Combined analysis of experiment 1 and 2: effectiveness and equity

Perceived fairness of payment schemes and conservation effort have a positive association. First, looking at the aggregated results from the previous subsections, the environmental effectiveness of treatments (Fig. 4) and their perceived equity (Fig. 5) follow the same ranking (although differences between payment schemes are larger in Experiment 2). For example, the *merit input* treatment induces the highest conservation effort in Experiment 1 and is also perceived as the most equitable in Experiment 2.

In addition, we investigate the correlation on the level of the individual land user. From each individual’s performance, we subtract the mean performance in his or her village to obtain an adjusted effort value (i.e., mean-centering, a form of adjusting for village effects). The correlation of adjusted conservation effort under a payment design and the equity score of the respective payment design is positive (Pearson’s $r = 0.19$, $p = 0.010$), as shown in Fig. 6.[9] Note that, to a large degree, this correlation is driven by the high environmental effectiveness and equity of differentiated payments conditional on input (*merit input*; orange dots), as compared to differentiated unconditional payments (*fixed individual*; blue dots). There is a negligible difference in effectiveness between input- and output-based conditional payments (Fig. 4), but a large difference in equity scores (Fig. 5). This indicates that the empirical association we observe between equity and

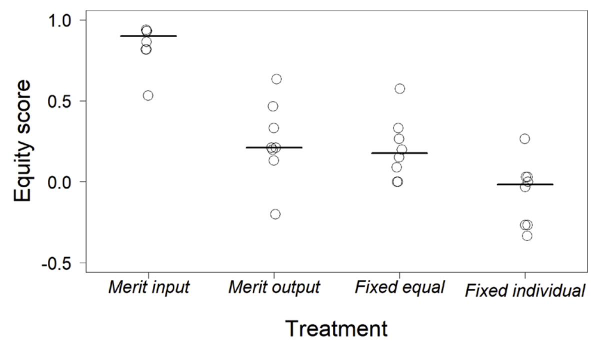

Fig. 5. Equity scores from the coordination game (Experiment 2). One equity score on the village level was derived for all treatments (i.e., payment designs; for explanations, see Table 1) in each village (n = 8). Each point is one observation (one village). Horizontal lines indicate medians.

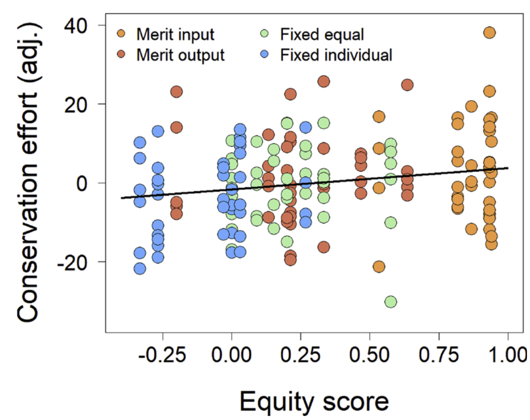

Fig. 6. Scatter plot of equity score on the village level (Experiment 2) and adjusted individual-level conservation effort (Experiment 1), aggregated over treatments. Adjusted conservation effort is calculated as difference from the village mean to control for the village effect. Line derived from linear regression (for interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

effectiveness across payment schemes might be positive overall but should not be generalized as being linear.

5. Discussion

In this study, we experimentally tested the environmental effectiveness (motivation of conservation effort) and social equity (collectively perceived fairness) of popular design features like payment differentiation and payment conditionality in incentive-based conservation policy. Outcomes were elicited in a real-effort task with real environmental benefits and a coordination game, respectively. Among 259 small-scale land users from North-Western Vietnam who participated in the incentivized lab-in-the-field experiments, we found that equity and effectiveness differed significantly across payment schemes. In both outcome dimensions, merit-based payments which were conditional on *input*, i.e. compensation for effort or costs were the most preferable payment scheme. The effectiveness differences, however, were rather small. Thus, we found that payments which are built on meritocratic principles (i.e., giving more to ecosystem service providers who – other things being equal – invest more time, money or physical effort) do not necessarily trade off effectiveness against equity, but can achieve both. Indeed, we found a weak positive correlation of effectiveness and equity across all four payment schemes that were employed in the experiments. In what follows, we discuss the results in more detail, along their implications and limitations.

The effectiveness premium of conditional payments based on *input*

[7] Simulations show that, given our group sizes of ten and eleven in Experiment 2, random responses would lead to roughly one half of subjects matching their village’s modal response(s). The proportion we find is substantially higher (65.4%). The simulated value approaches 25% for larger group sizes. We thank Carl Salk for pointing this out.

[8] These numbers should be interpreted with caution, as the sample size for this comparison is small (n = 8 at the village level) which implies a high risk of a type II error (false negative).

[9] The value is lower without adjustment (Pearson’s $r = 0.14$, $p = 0.076$). A linear fixed-effect model yields similar results to our adjusted performance correlation (β: 5.69 ± 2.17, $p = 0.010$). Yet, we decided to use correlation coefficients here to emphasize that one cannot infer any causal relationships from the combined data of both experiments.

(compensation for effort or costs) was particularly large when they were compared to unconditional payments that differentiated payments between recipients (treatment *fixed individual*; $p < 0.05$). The gain was slightly smaller (and missed statistical significance at $p = 0.14$) when they were compared to unconditional payments that were undifferentiated (i.e., egalitarian payments; treatment *fixed equal*). Thus, when focusing only on the effectiveness for motivating conservation behavior, unconditional egalitarian payments might be broadly comparable to merit-based payments, at least for the effect sizes which we were able to detect given the statistical power of the study. On the other hand, unconditional payments which are *not* egalitarian (i.e., which differentiate) appear clearly less effective than merit-based payments. The latter case is potentially more common in real-life PES implementations than intended by policy-makers: Cases where elite capture, gender inequality, corruption or opaque and complex benefit distribution mechanisms exclude portions of the population from financial transfers create payment differentiation without conditioning on conservation effort. Our results suggest that such *de facto* arbitrary modes of differentiation inhibit environmental effectiveness (Experiment 1) and, further, are seen as most unfair and socially inappropriate (Experiment 2).

The results on the effectiveness of incentives for conservation effort conditional on individual merit (Experiment 1) are broadly in line with previous lab-in-the-field experiments where individual incentives worked better than group incentives (Gatiso et al., 2018) and where conditionality on a measure of *change* in ecosystem services provision rather than absolute ecosystem services provision (reflecting more control over the outcome by ecosystem services providers; Reutemann et al., 2016) led to more forest-friendly behavior. Such payment differentiation conditional on ecosystem services providers' individual costs is indeed seen as a desirable design feature of PES which, however, is globally not (yet) widely practiced (Engel, 2016; Wunder et al., 2018), despite its apparent contribution to positive ecological outcomes of PES (Ezzine-de-Blas et al., 2016).

The case our study makes for these payment design features becomes stronger when considering that input-based conditional payments were also perceived as more equitable than alternative payment designs in the coordination games (Experiment 2). Given the, at least in practice, rather unconditional (with respect to individual conservation costs and efforts) and undifferentiated distribution of benefits in the Vietnamese PFES program (Loft et al., 2017b), our results cast some doubt on its long-term ecological and social sustainability. Not only are the majority of the formally determined payment indicators (e.g., origin of forest) only poor indicators of actual conservation effort and costs by land users (see section 2), but often payments are distributed fully undifferentiated due to practical limitations. Based on our results, we can conclude that payment schemes without conditionality on individual conservation effort and with differentiation according to factors out of the control of local ecosystem services providers, are perceived as socially inappropriate or unfair. Even when ignoring consequences for effectiveness, this corroborates the conclusion that current PFES legislation interferes with social norms of distributive fairness. However, local actors, their social and cultural norms play an important role in adapting and transforming PES institutions. Ignoring their notions and excluding them from decision-making may reinforce existing inequity (Van Hecken et al., 2015).

Importantly, our study ignores some obstacles for the implementation of well-differentiated, conditional payments like a lack of perfect information on individual efforts and the high administrative costs that would be associated with gathering such information (Engel, 2016). Thus, one also must consider the efficiency of alternative PES schemes, a dimension that our study does not explicitly address. Arguably, the rather small gain in effectiveness by merit-based payments observed in the real-effort task, i.e. 9% increase in conservation effort in merit-based payments conditional on input vs. unconditional egalitarian payments (Table 3) might not justify an implementation solely based on effectiveness criteria. This is because costs must eventually be weighed against benefits in policy-making. While conditional payments that reward merit, and in particular those that purely reward effort, may induce slightly higher conservation efforts, they could also imply higher bureaucracy and monitoring costs. A resulting lack of transparent and reliable monitoring, however, could in turn spur perceptions of unfairness, even under schemes that are intended to be conditional and well-differentiated. It is an important empirical question to determine how much external monitoring is necessary for an implementation of payments that are differentiated according to conservation effort. Interestingly, Jindal et al. (2017) find that, in a sample of land users in Northern Vietnam which is similar to ours, subjects "report truthfully", i.e., they collect the rewards they have earned in a real effort task *but not more*, even though they could easily cheat without being detected. The authors suggest that honesty norms and trust can enable self-monitoring. Similarly, other lab-in-the-field experiments showed that, even when strictly monitored PES is abandoned, land users might carry on norms of pro-conservation behavior (Andersson et al., 2018).[10] Thus, in some societies, mutual monitoring and internal benefit distribution could be a feasible alternative to external monitoring, lowering risks and costs for the implementation of differentiated and conditional PES (Engel, 2016).

Note that sensitivity for monetary rewards was generally small in our real-effort task (Fig. 4). The unconditional payments in two of our treatments would have implied full rewards even in the absence of any effort, but still they led to high conservation effort. This finding is in line with a potentially important role of environmental motives and intrinsic motivation in conservation effort (Rode et al., 2015). In the survey after the experiment, 65% of participants stated that they were motivated to exert effort in the task, because, amongst other things, they wanted to support conservation. An experimental design which does not produce real-world environmental public goods, such as standard social dilemma games, could have missed this important motive, arguably with reductions in external validity. The surprisingly strong deviation from the "rational" prediction of no effort in unconditional payment schemes mirrors findings from other real-effort experiments in behavioral economics under more controlled conditions (Araujo et al., 2016; Eckartz, 2014). Explanations could be competitiveness or conformity. It could also be that subjects want to please the experimenters (Zizzo, 2010). Although subjects were successfully encouraged not to communicate verbally with each other in our setup, they could in principle observe the performance of their neighbors, which might induce peer pressure. A general curiosity for the task which could have played a role, too. Increasing the available time for the task may be one way to address this problem (Charness et al., 2018).

In addition to the effects of payment design, we found a comparably large and significant gender effect, with women exerting greater efforts. The effect remains after controlling for experience with the task and other potential confounders. Further, it is an individual-level effect: the share of women in a village session is not correlated with performance on the session level (Supplementary material). In our real effort task, the packing of soil bags is regarded as a tedious and repetitive activity, often conducted by women in Vietnamese nurseries and forest enterprises (VAN 2015). In many parts of Vietnam, soil bag packing for afforestation is also considered a poverty reduction strategy to enhance gender equality by supporting young women (Nghe An Provincial Party Committee 2016). Moreover, due to social and gender norms, packing soil bags might be seen as rather suitable for women. Thus, on the one hand, the effect could be interpreted as reflecting that our task is culturally rendered as rather "female." On the other hand, it gives

[10] Although abandoning PES is different from abandoning external monitoring, both require community norms to replace an external authority – a leviathan.

additional confidence for the conjecture that increasing women participation in forest management and conservation, e.g. in afforestation-related activities, will be beneficial in terms not only of gender equality, but also environmental effectiveness (Cook et al., 2019). Other PFES studies in Vietnam (Haas et al., 2019; Pham et al., 2018) found that there is inequality in PFES participation, as only strong and young male villagers are selected for forest patrolling groups, and therefore these men are the main PFES beneficiaries. Our findings call for a gender-sensitive approach in PFES management where roles of both, women and men, are recognized and payment distributions are developed based on these gender role differences. Many development interventions like microfinance or conditional cash transfers already target women as the primary agents of change (Duflo, 2012; Ma et al., 2017).

In conclusion, we showed that equity and effectiveness of PES designs are affected by the differentiation and conditionality of payments and that equitable PES designs can coincide with effective PES designs. There is a tendency that both are maximized when individuals receive differentiated payments which are conditional on costs of and effort for ES provision, although, regarding effectiveness, effect estimates are small and noisy. Wunder et al. (2018, p. 149) state that "navigating efficiency and equity trade-offs in the face of contextual fairness principles can lead to some hard choices for PES design." Our study, however, offers the promising conclusion that this might not always be the case. One of the big advantages of our experimental set-up is that we can separately measure effectiveness and equity in the same population at the same time and based on the same type of conservation intervention (paying individuals for filling soil bags for afforestation). However, we are not able to make claims about causal links between equity and effectiveness. While our results are consistent with the view that there is an "instrumental value" of equity for the effectiveness of conservation policies (Law et al., 2018), the research design does not allow a causal attribution of any differences in effort to perceived fairness of the payment scheme. This is an important question for future research and we think that the experimental paradigm presented here could help to further investigate this, because it constitutes a middle-ground between large-scale experimentation with actual policy interventions (e.g., Randomized Control Trials, Jayachandran et al., 2017) and abstract lab-in the-field experiments based merely on social dilemma games. Arguably, shortcomings of the presented design in its current form are that many potential confounders, e.g., peer pressure due to observability of effort and the uncertainty about its generalizability to other contexts and behaviors, such as patrolling in the forest remain. Note that we also have only looked at aspects of distributional equity, although equity is a multidimensional concept (Myers et al., 2018) and can be defined along different dimensions in PES (Brown and Corbera, 2003; Corbera et al., 2007; Friedman et al., 2018; Martin et al., 2014a, 2014b; McDermott et al., 2013; Pascual et al., 2010, 2014). Contextual and procedural equity, as well as recognition, might be equally important (Pascual et al., 2014). Future experiments could also investigate effort under the premise that people can self-select into different payment schemes vs. an exogenous subscription (Rommel, 2015). This would allow us to estimate the impact of "letting the people have a say in how they want to be paid" on effort. Thus, our novel experimental designs offers a number of interesting routes to further investigate in how far exactly in conservation policies "equity and efficiency are fundamentally codependent" (Law et al., 2018, p. 300).

Acknowldegements

We thank Prof. Dr. K. Müller for making this study possible. We are grateful for the field support by the Dien Bien Agriculture Extension Center, in particular by Mrs. Dinh Thi Thu Ha and her team. The manuscript was greatly improved by comments from Carl Salk and the anonymous reviewers. Feedback on the study at the FLARE Forests & Livelihoods Network Meeting 2017, the Lund Conference on Earth System Governance 2017, the 2nd meeting of the Research Network on Economic Experiments for the Common Agricultural Policy, and the 4th Workshop on Experimental Economics for the Environment, 2019 further helped to improve the manuscript. The drawing of the real-effort task was designed by Felix Korda.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.landusepol.2019.05.010.

References

Andersson, K.P., Cook, N.J., Grillos, T., Lopez, M.C., Salk, C.F., Wright, G.D., Mwangi, E., 2018. Experimental evidence on payments for forest commons conservation. Nat. Sustainability 1, 128.
Handberg, Ø.N., Angelsen, A., 2019. Pay little, get little; pay more, get a little more: a framed forest experiment in Tanzania. Ecol. Econ. 156, 454–467.
Araujo, F.A., Carbone, E., Conell-Price, L., Dunietz, M.W., Jaroszewicz, A., Landsman, R., Lamé, D., Vesterlund, L., Wang, S.W., Wilson, A.J., 2016. The slider task: an example of restricted inference on incentive effects. J. Econ. Sci. Assoc. 2, 1–12. https://doi.org/10.1007/s40881-016-0025-7.
Ariely, D., Kamenica, E., Prelec, D., 2008. Man's search for meaning: the case of legos. J. Econ. Behav. Organ. 67, 671–677. https://doi.org/10.1016/j.jebo.2008.01.004.
Bennett, N.J., 2016. Using perceptions as evidence to improve conservation and environmental management: perceptions and conservation. Conserv. Biol. 30, 582–592. https://doi.org/10.1111/cobi.12681.
Bolton, G.E., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. Am. Econ. Rev. 90, 166–193.
Brown, K., Corbera, E., 2003. Exploring equity and sustainable development in the new carbon economy. Clim. Policy 3, S41–S56.
Calvet-Mir, L., Corbera, E., Martin, A., Fisher, J., Gross-Camp, N., 2015. Payments for ecosystem services in the tropics: a closer look at effectiveness and equity. Curr. Opin. Environ. Sustainability 14, 150–162.
Charness, G., Gneezy, U., Henderson, A., 2018. Experimental methods: measuring effort in economics experiments. J. Econ. Behav. Organ. 149, 74–87.
Chu, L., Grafton, R.Q., Keenan, R., 2019. Increasing conservation efficiency while maintaining distributive goals with the payment for environmental services. Ecol. Econ. 156, 202–210.
Cook, N.J., Grillos, T., Andersson, K.P., 2019. Gender quotas increase the equality and effectiveness of climate policy interventions. Nat. Clim. Change 1.
Corbera, E., Kosoy, N., Tuna, M.M., 2007. Equity implications of marketing ecosystem services in protected areas and rural communities: case studies from Meso-America. Global Environ. Change 17, 365–380.
DellaVigna, S., Pope, D., 2018. What motivates effort? Evidence and expert forecasts. Rev. Econ. Stud. 85, 1029–1069. https://doi.org/10.1093/restud/rdx033.
Duflo, E., 2012. Women empowerment and economic development. J. Econ. Lit. 50, 1051–1079. https://doi.org/10.1257/jel.50.4.1051.
Eckartz, K., 2014. Task Enjoyment and Opportunity Costs in the Lab: The Effect of Financial Incentives on Performance in Real Effort Tasks (No. 2014-005). Jena Economic Research Papers.
Engel, S., 2016. The devil in the detail: a practical Guide on designing payments for environmental services. Int. Rev. Environ. Res. Econ. 9, 131–177. https://doi.org/10.1561/101.00000076.
Erkal, N., Gangadharan, L., Koh, B.H., 2018. Monetary and non-monetary incentives in real-effort tournaments. Eur. Econ. Rev. 101, 528–545.
Ezzine-de-Blas, D., Wunder, S., Ruiz-Pérez, M., Moreno-Sanchez, R., del, P., 2016. Global patterns in the implementation of payments for environmental services. PLoS One 11, e0149847. https://doi.org/10.1371/journal.pone.0149847.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 817–868.
Friedman, R.S., Law, E.A., Bennett, N.J., Ives, C.D., Thorn, J.P.R., Wilson, K.A., 2018. How just and just how? A systematic review of social equity in conservation research. Environ. Res. Lett. 13, 053001. https://doi.org/10.1088/1748-9326/aabcde.
Gatiso, T.T., Vollan, B., Vimal, R., Kühl, H.S., 2018. If possible, incentivize individuals not groups: evidence from lab-in-the-field experiments on forest conservation in rural Uganda. Conserv. Lett. 11, e12387.
Gehrig, S., Schlüter, A., Hammerstein, P., 2019. Sociocultural heterogeneity in a common pool resource dilemma. PLoS One 14, 1–16. https://doi.org/10.1371/journal.pone.0210561.
Gintis, H., 2009. The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences. Princeton University Press.
Haas, J.C., Loft, L., Pham, T.T., 2019. How fair can incentive-based conservation get? The interdependence of distributional and contextual equity in vietnam´s payments for Forest environmental services program. Ecol. Econ. 160, 205–2014.
Halpern, B.S., Klein, C.J., Brown, C.J., Beger, M., Grantham, H.S., Mangubhai, S., Ruckelshaus, M., Tulloch, V.J., Watts, M., White, C., et al., 2013. Achieving the triple bottom line in the face of inherent trade-offs among social equity, economic return, and conservation. Proc. Natl. Acad. Sci. 110, 6229–6234.
Hayes, T., Murtinho, F., 2018. Communal governance, equity and payment for ecosystem services. Land Use Policy 79, 123–136. https://doi.org/10.1016/j.landusepol.2018.08.001.

He, J., Sikor, T., 2015. Notions of justice in payments for ecosystem services: insights from China's sloping Land conversion program in Yunnan Province. Land Use Policy 43, 207–216. https://doi.org/10.1016/j.landusepol.2014.11.011.

Jayachandran, S., De Laat, J., Lambin, E.F., Stanton, C.Y., Audy, R., Thomas, N.E., 2017. Cash for carbon: a randomized trial of payments for ecosystem services to reduce deforestation. Science 357, 267–273.

Jindal, R., Arvai, J., Catacutan, D., Bac, D.V., et al., 2017. To cheat or not? Results from behavioral experiments on self-monitoring in Vietnam. Strategic Behav. Environ. 7, 179–219.

Kerr, J., Vardhan, M., Jindal, R., 2012. Prosocial behavior and incentives: evidence from field experiments in rural Mexico and Tanzania. Ecol. Econ. 73, 220–227.

Konow, J., 2003. Which is the fairest one of all? A positive analysis of justice theories. J. Econ. Lit. 41, 1188–1239.

Kosfeld, M., Neckermann, S., Yang, X., 2017. The effects of financial and recognition incentives across work contexts: the role of meaning. Econ. Inq. 55, 237–247. https://doi.org/10.1111/ecin.12350.

Krupka, E.L., Weber, R.A., 2013. Identifying social norms using coordination games: why does dictator game sharing vary? J. Eur. Econ. Assoc. 11, 495–524.

Law, E.A., Bennett, N.J., Ives, C.D., Friedman, R., Davis, K.J., Archibald, C., Wilson, K.A., 2018. Equity trade-offs in conservation decision making: equity in decision making. Conserv. Biol. 32, 294–303. https://doi.org/10.1111/cobi.13008.

Leimona, B., Carrasco, L.R., 2017. Auction winning, social dynamics and non-compliance in a payment for ecosystem services scheme in Indonesia. Land Use Policy 63, 632–644. https://doi.org/10.1016/j.landusepol.2015.10.022.

Loft, L., Pham, T.T., Wong, G.Y., Brockhaus, M., Le, D.N., Tjajadi, J.S., Luttrell, C., 2017a. Risks to REDD+: potential pitfalls for policy design and implementation. Environ. Conserv. 44, 44–55.

Loft, L., Le, D.N., Pham, T.T., Yang, A.L., Tjajadi, J.S., Wong, G.Y., 2017b. Whose equity matters? National to local equity perceptions in Vietnam's payments for Forest ecosystem services Scheme. Ecol. Econ. 135, 164–175.

Luttrell, C., Loft, L., Gebara, M.F., Kweka, D., Brockhaus, M., Angelsen, A., Sunderlin, W.D., 2013. Who should benefit from REDD+? Rationales and realities. Ecol. Soc. 18.

Ma, Z., Bauchet, J., Steele, D., Godoy, R., Radel, C., Zanotti, L., 2017. Comparison of direct transfers for human capital development and environmental conservation. World Dev. 99, 498–517. https://doi.org/10.1016/j.worlddev.2017.05.030.

Martin, A., Gross-Camp, N., Kebede, B., McGuire, S., 2014a. Measuring effectiveness, efficiency and equity in an experimental payments for ecosystem services trial. Global Environ. Change 28, 216–226.

Martin, A., Gross-Camp, N., Kebede, B., McGuire, S., Munyarukaza, J., 2014b. Whose environmental justice? Exploring local and global perspectives in a payments for ecosystem services scheme in Rwanda. Geoforum 54, 167–177.

McDermott, M., Mahanty, S., Schreckenberg, K., 2013. Examining equity: a multidimensional framework for assessing equity in payments for ecosystem services. Environ. Sci. Policy 33, 416–427.

McElwee, P., Nghiem, T., Le, H., Vu, H., Tran, N., 2014. Payments for environmental services and contested neoliberalisation in developing countries: a case study from Vietnam. J. Rural Stud. 36, 423–440.

Muradian, R., Corbera, E., Pascual, U., Kosoy, N., May, P.H., 2010. Reconciling theory and practice: an alternative conceptual framework for understanding payments for environmental services. Ecol. Econ. 69, 1202–1208.

Myers, R., Larson, A.M., Ravikumar, A., Kowler, L.F., Yang, A., Trench, T., 2018. Messiness of forest governance: how technical approaches suppress politics in REDD+ and conservation projects. Global Environ. Change 50, 314–324.

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

Narloch, U., Pascual, U., Drucker, A.G., 2013. How to achieve fairness in payments for ecosystem services? Insights from agrobiodiversity conservation auctions. Land Use Policy 35, 107–118. https://doi.org/10.1016/j.landusepol.2013.05.002.

Nelson, K.M., Schlüter, A., Vance, C., 2018. Funding conservation locally: insights from behavioral experiments in Indonesia. Conserv. Lett. 11, e12378.

Pascual, U., Muradian, R., Rodríguez, L.C., Duraiappah, A., 2010. Exploring the links between equity and efficiency in payments for environmental services: a conceptual approach. Ecol. Econ. 69, 1237–1244. https://doi.org/10.1016/j.ecolecon.2009.11.004.

Pascual, U., Phelps, J., Garmendia, E., Brown, K., Corbera, E., Martin, A., Gomez-Baggethun, E., Muradian, R., 2014. Social equity matters in payments for ecosystem services. Bioscience 64, 1027–1036.

Pham, T., Bennett, K., Vu, T.P., Brunner, J., Le Ngoc, D., Nguyen, D.T., et al., 2013. Payments for Forest Environmental Services in Vietnam: From Policy to Practice. CIFOR Occasional Paper.

Pham, T., Moeliono, M., Wong, G.Y., Brockhaus, M., Dung, L.N., 2018. The politics of swidden: a case study from nghe an and Son La in Vietnam. Land Use Policy. https://doi.org/10.1016/j.landusepol.2017.10.057.

Reutemann, T., Engel, S., Pareja, E., 2016. How (not) to pay—field experimental evidence on the design of REDD+ payments. Ecol. Econ. 129, 220–229.

Rode, J., Gómez-Baggethun, E., Krause, T., 2015. Motivation crowding by economic incentives in conservation policy: a review of the empirical evidence. Ecol. Econ. 117, 270–282. https://doi.org/10.1016/j.ecolecon.2014.11.019.

Rommel, J., 2015. What can economic experiments tell us about institutional change in social–ecological systems? Environ. Sci. Policy 53, 96–104. https://doi.org/10.1016/j.envsci.2014.05.006.

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. Conserv. Lett. 10, 414–421.

Sattler, C., Matzdorf, B., 2013. PES in a nutshell: from definitions and origins to PES in practice—Approaches, design process and innovative aspects. Ecosyst. Serv. 6, 2–11. https://doi.org/10.1016/j.ecoser.2013.09.009.

Schomers, S., Matzdorf, B., 2013. Payments for ecosystem services: a review and comparison of developing and industrialized countries. Ecosyst. Serv. 6, 16–30.

Sills, E.O., de Sassi, C., Jagger, P., Lawlor, K., Miteva, D.A., Pattanayak, S.K., Sunderlin, W.D., 2017. Building the evidence base for REDD+: study design and methods for evaluating the impacts of conservation interventions on local well-being. Global Environ. Change 43, 148–160.

Torres-Guevara, L.E., Schlüter, A., 2016. External validity of artefactual field experiments: a study on cooperation, impatience and sustainability in an artisanal fishery in Colombia. Ecol. Econ. 128, 187–201. https://doi.org/10.1016/j.ecolecon.2016.04.022.

Van Hecken, G., Bastiaensen, J., Windey, C., 2015. Towards a power-sensitive and socially-informed analysis of payments for ecosystem services (PES): addressing the gaps in the current debate. Ecol. Econ. 120, 117–125. https://doi.org/10.1016/j.ecolecon.2015.10.012.

Vatn, A., 2010. An institutional analysis of payments for environmental services. Ecol. Econ. 69, 1245–1252.

Vatn, A., 2015. Markets in environmental governance. From theory to practice. Ecol. Econ. 117, 225–233.

Vietnam Forest Protection and Development Fund (VNFF), 2015. Report on Three Years of Implementation of Policy on Payment for Forest Environmental Services in Viet Nam (2011–2014).

Wunder, S., 2005. Payments for environmental services: some nuts and bolts. CIFOR Occasional Paper 42, 24.

Wunder, S., 2015. Revisiting the concept of payments for environmental services. Ecol. Econ. 117, 234–243.

Wunder, S., Brouwer, R., Engel, S., Ezzine-de-Blas, D., Muradian, R., Pascual, U., Pinto, R., 2018. From principles to practice in paying for nature's services. Nat. Sustainability 1, 145.

Zizzo, D.J., 2010. Experimenter demand effects in economic experiments. Exp. Econ. 13, 75–98. https://doi.org/10.1007/s10683-009-9230-z.