# Supplementary Information for 'Fair payments for effective environmental conservation'

Lasse Loft, Stefan Gehrig, Carl Salk, Jens Rommel

## Contents

# 1 Supplementary Methods

## 1.1 Experimental design and real-world conservation incentives

Real-effort experiments are a common method in experimental economic research on the role of incentives (1). They have recently gained popularity in conservation research with natural resource users in the field (e.g., (2)). Acknowledging that economic experiments are abstractions of real-world settings, our aim was to design a task that reflects the real-world setting as much as possible (external validity) while remaining experimentally clean and tractable to allow causal inference (internal validity). Here, we explain how key building blocks of the experimental design – the choice of task (see 1.1.1) and treatments (see 1.1.2) – resemble features of real-world environmental incentive schemes. Our design aims to test the effect of (in)equity in incentive-based conservation, i.e., the instrumental role that equity can play in shaping outcomes (3).

### 1.1.1 Choice of task

Incentives are often provided for the conservation of natural habitats or the generation of "environmental amenities that are public goods" ((4), p. 247). A key step in our research was to choose a task that resembles conservation instruments that provide payments for conducting activities which improve the state of natural resources (5). We argue that an experiment which measures real effort for conservation is suited for research on incentive-based conservation generally (6). Globally, there are many examples of payment schemes that incentivize activities such as tree planting (7). In Vietnam's national Payments for Forest Ecosystem Services program, "participants are not being paid based on outcome in terms of a measured quantity of ES. Rather, the payment is based on labour input that should lead to an increased provision of ecosystem services", i.e. conducting forest protection activities, such as forest patrolling, reporting of illegal timber logging and hunting, and fighting of forest fires ((8), p. 206), i.e., conducting forest protection activities, such as forest patrolling, reporting illegal timber logging and hunting, and fighting forest fires (9). China's Sloping Land Conversion Program (SLCP) provides farmers with in-kind and cash payments in exchange for converting cropland on steep slopes to less erodible forest and grassland (10). The payments are linked to allowing forest regeneration, aerial seeding and artificial planting (11). In Mexico's Federal Conservation Payments Program, incentivized activities include reforestation (12), maintenance of existing forest or natural land cover and engagement in land management activities such as "building fences, controlling pests, or patrolling for illegal activity"((5), p. 7016)). In Europe, Agri-Environmental Measures are incentive-based instruments that provide compensation payments to farmers in return for the temporary adoption of, for example, landscape and habitat measures such as planting wild-

flower strips on the margins of plots in a wide range of agricultural landscapes ((13), (14), (15)). Many afforestation and reforestation projects under the UNFCCC Kyoto Protocol Clean Development Mechanism paid for tree-planting activities in developing countries (16).

For the real-effort experiment, we needed a task that reflects real-world conservation activities, makes effort easily quantifiable and requires rather simple physical effort for which skill or ability would not vary greatly among subjects. Further, the task should produce real-world environmental benefits which distinguishes it from abstract pen-and-paper experiments on payments for ecosystem services in which subjects invest in or harvest hypothetical resources ((17), (18), (19)). We thus chose filling biodegradable bags with fertile soil to grow seedlings for reforestation as an activity that meets all of these requirements. Producing seedlings for reforestation is an activity that is either part of many of the above mentioned conservation programs or an activity directly supporting these actions. In the study region, soil bags are commonly produced and planted by conservation NGOs and tree nurseries. Just as in our task, organizations pay people for filling bags which are then used for afforestation. In our study, 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 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 other factors like payments. Lab-in-the-field experiments that produce real-world public goods are only rarely employed (but see (20) and (21) for experiments that fund schools and bridges, respectivaly), in particular in combination with real-effort tasks (but see (22)). We have employed a similar real-effort task in a previous study of a different topic, a comparison of equity preferences and effort effects between different payment distribution mechanisms (6).

### 1.1.2 Choice of treatments

Besides the task, it was important to design treatments so that they capture the problem of inequities in incentive-based conservation. The randomly-assigned disparity in payment rates within the same group of subjects (high rate vs low rate) needs to reflect payment inequalities found in many real-world programs leading to unjustified or poorly-justified income disparities.

In Vietnam's national Payments for Forest Ecosystem Services program, payment distribution to households is not clearly specified in the underlying legislation (23). Much is left to be decided on the local level, and "officials have much latitude in recognizing household ownership and access over forests and distributing these funds among and within villages" (9). Previous research found that people are paid differently for the same conservation activities like forest patrolling, reporting illegal timber logging or hunting activities and fighting forest fires. One

reason for this is that the amount of funds that any commune receives depends on its forest area, and payment levels per household are therefore related to the ratio of forest area to population in a given commune (9). A second reason is that payment rate calculation considers the type of forest (whether it is special-use, protection or production forest), and the origin of forest (whether it is natural or planted). Depending on the forest classification, different payments are being delivered, while work effort for patrolling these forests is the same (8). Overall, "the result is that the level of payments to participating households can vary significantly, even between neighbouring communes or even within a village" ((9), p. 265-266). While payment inequality on very local scales is seen in Vietnam, less extreme examples are common elsewhere. For instance, in the case of SLCP, different rates were paid for the same activities, depending on whether the activity was being conducted in the Yangtze River Basin or the upper and middle reaches of the Yellow River Basin (11). He and Sikor have explained this difference by noting different opportunity costs in these areas (24). Further, the duration of the payments differed according to which type of forest was planted. Payments were delivered for 8 years if "ecological forests" were planted and for 5 years or 2 years if "economic forests" or grasses were planted (25). While these payment differences were made transparent on paper and justified by different opportunity costs, early studies on SLCP implementation find evidence that there were systematic differences in local implementation. Uchida et al. (26) state that "the actual delivery of payments often falls short (…) in some villages, (…)" (p. 600). Among the listed reasons for full compensation not reaching participating farmers are that "village officials may deduct labor cost to plant trees on the farmer's program plots. They may deduct back-taxes owed by the farmer. In other cases, some plots are retired before being fully certified under the program". Bennett (25) found that about two-thirds of the 360 surveyed households could not "choose what kinds of trees to plant on their enrolled land" (p. 706-707). Thus, since the total payment depended on the type of forest planted, income differences resulted at the village level, which the villagers had no control over.

The occurrence of payment inequality in the cases of Vietnam and China are only partially due to unintended consequences of program designs; payment differences within the same scheme are to some degree desired and justified. Yet, during practical implementation on the ground, unjustified payment inequality often occurs even when it has not been intended by legislation (see (3) for additional examples of inequity in incentive-based conservation). Payment inequality in incentive-based conservation is not restricted to the context of developing countries and economies in transition, but commonplace in industrialized countries too. In The European Union's Agri-Environmental Measures scheme, "payment levels can vary substantially from one Member State to another for a range of reasons including differences between sites and the fact that not all Member States pay the full amount of income foregone and costs incurred" ((27), p. 10). Within EU member states that have a federal structure, such as Germany, planting of

wildflower strips is financially compensated differently depending on the total budget available in the federal state. As a consequence, neighboring farms (in different states) can receive different payments for the same activity.

## 1.2 Data collection

### 1.2.1 Village selection process

Our study region was Nam Dong district in Thừa Thiên Huế Province in Central Vietnam. The province is part of the Central Annamite Ecoregion, a global biodiversity hotspot. The district was selected, because it is located in the buffer zone of Bạch Mã National Park, the center of the last remaining rain forest in Central Vietnam. The area has been threatened by deforestation and defaunation (in the survey, 32% and 30% of subjects responded that they know people in their village who go logging or hunting in the forest, respectively) and is an important target for national and international biodiversity conservation initiatives. For the selection of experimental sites, we collected socio-demographic background data from 60 villages in the 11 communes of Nam Dong district (see 1.5). These villages were filtered according to the criteria of having more than 250 inhabitants and having forest land enrolled in the national Payments for Forest Ecosystem Services (PFES) progam. From the resulting subset (N = 33), 15 villages were randomly selected. One village served as a pilot to test that the task is feasible and understood, calibrate payments and train staff with a smaller sample (N = 14 subjects) and a single treatment only (i.e., we did *not* collect data on effects of treatment manipulations in the pilot). Consequently, we conducted 14 full sessions in 14 villages. In cases where sampled villages were geographically close, we scheduled sessions for the same day or consecutive day when possible to avoid contamination and spillover effects between villages. To meet the minimum pre-specified sample size of N = 432 (see 1.4) and have equally-sized treatment groups of eight in each village, we recruited 32 participants in each of the 14 villages.

### 1.2.2 Schedule of experiments

In total, the visit to a village, including the preparation of the task and survey, the debriefing, and the clean-up took about 3.5 hours on average. We conducted morning and afternoon sessions. The morning sessions started at around 8 am, the afternoon sessions at around 2 pm. Table S1 reports date and time for all villages where the experiment was conducted.

**Table S1:** Sampling schedule for the 14 experimental sites and one pilot site (first row).

| Village | Commune | Date | Time |
|---|---|---|---|
| 8 | Huong Hoa | 2019-03-24 | Morning |
| Cha Mang | Thuong Lo | 2019-03-26 | Afternoon |
| 6 | Thuong Long | 2019-03-27 | Morning |
| 8 | Thuong Long | 2019-03-27 | Afternoon |
| La Van | Thuong Nhat | 2019-03-28 | Morning |
| Ta Lu | Thuong Nhat | 2019-03-28 | Afternoon |
| Ria Ho | Thuong Lo | 2019-03-29 | Morning |
| Ta Rung | Huong Son | 2019-03-30 | Morning |
| Bha Bha | Huong Son | 2019-03-30 | Afternoon |
| 1 | Thuong Quang | 2019-03-31 | Morning |
| 5 | Thuong Quang | 2019-03-31 | Afternoon |
| 6 | Thuong Quang | 2019-04-01 | Morning |
| Ta Rinh | Thuong Nhat | 2019-04-01 | Afternoon |
| A2 | Huong Son | 2019-04-02 | Morning |
| Pa Noong | Huong Son | 2019-04-02 | Afternoon |

### 1.2.3 Enumerators

Six enumerators, one group facilitator, and one driver were subcontracted and selected by our local implementation partner WWF Vietnam. This team accompanied the principal research team in the field (L.L., J.R., and one research assistant). The enumerators spoke English and Vietnamese. All of them had a background in forestry and/or resource economics and most of them had experience working in the field. Prior to visiting the study villages, a four-day training was conducted with enumerators. One village visit was used to pilot the procedure. Afterwards, a debriefing took place with enumerators, the procedure was revised, and additional training and adjustments of field procedures took place.

### 1.2.4 Collaboration with village heads

All Vietnamese villages have an elected official, the so-called village head whom we used as entry points to the community. This had the advantage that a sample of 32 members of a village was readily available for participation in each village. Village heads were called by the local assistants three days prior to conducting the experiments. Village heads were informed that an international team of researchers would conduct an exercise related to reforestation activities and a brief survey. They were not informed about details or research questions. Village heads were instructed that the activity would require:

- 32 participants
- place(s) in the village were 32 participants could meet and work with sufficient space;
- participants would get paid for the work and the participation in the survey;
- two villagers would need to assist the research team with the setup of soil piles, which would be delivered the day prior to the visit.

### 1.2.5 Task setup

On the day of the experiment, the soil preparation task was prepared for four groups of eight participants separately (see Figure 2 in main text for a schematic drawing and Figure S1 for examples of a group's task setup). Groups were located at some distance to each other to prevent communication or any other influences across treatments. Each group's task setup consisted of…

- …a circle of 8 plastic chairs for participants to sit on
- …eight boxes which fit 20 packed soil bags, one next to each chair
- …a pile of 0.5 $m^3$ of soil in the middle of the circle
- …a large clock visible to all participants of the group
- …a bucket of water and paper towel for washing hands
- …a table with water and snacks next to the circle

The areas where tasks were conducted were generally located in the shade to ensure similar working conditions for all groups.


**Figure S1:** Examples of task setups for two groups in two different villages.

### 1.2.6 Sampling of subjects in the village

Village heads informed villagers that 32 people would be able to participate in the task. A downside of this is that the selection process was out of our direct control. Note that randomization was unaffected by the selection process. Villagers were excluded if they were drunk, underage, or could not commit themselves to finish the task and the survey.

### 1.2.7 Experimental procedure

After arriving at the village, the group facilitator got in contact with the village head who gathered the participants. Then the facilitator introduced the research team and the enumerators to the participants. This followed a standardized format (see 1.2.10). After the introduction, an overview of the planned task was provided, it was explicitly stated that the task would be paid, and consent of participation was obtained (see 1.2.8). Consent forms were distributed and had to be signed by all participants. In case of illiteracy, the form was read aloud to participants, and the signature was substituted by a fingerprint. Note that 5 subjects withdrew their participation during the task and left the experiment early without survey and payment for performance, but received the show-up fee (for details, see Table S2). These subjects did

not comply with the protocol and were removed from the dataset for analysis, as specified in the pre-registration. No personal task or survey data for participants who withdrew from the experiment was retained.

**Table S2:** Description for participants who withdrew participation during the experiment.

| ID | Treatment | Comments on participant by enumerators |
|---|---|---|
| 33 | Unequal, High pay rate | left the activity after 30 minutes |
| 225 | Unequal, High pay rate | left to look for her child after 28 minutes |
| 281 | Equal, Low pay rate | stopped after 15 min and left for a wedding |
| 286 | Equal, Low pay rate | took 2 min to play with his mobile phone; stopped after 15 min and left for a wedding |
| 288 | Unequal, High pay rate | left for going home after 32 min because of her own business; was a health issue with her mother |

In the next step, detailed instructions on the soil bag preparation task were provided verbally in Vietnamese by the facilitator (see 1.2.10). After the demonstration of the task, each of the 32 participants drew a closed envelope with ID numbers from a pile of shuffled envelopes. Each of the four enumerators was randomly assigned to a group consisting of eight ID numbers (at this time being unaware of the group's treatment). The other two enumerators helped out with other tasks in the field, monitored rule enforcement and took notes of peculiarities among participants' behavior (e.g., exceptionally high or low work performance and the apparent reason for this, extensive snack consumption, extensive breaks taken, stated reasons for withdrwal from the experiment). Within the four separate groups, the procedure of soil bag packing was explained again. It was shown how to put the soil bags in the individual boxes and where to unload the individual boxes after 20 soil bags were reached. To help participants get familiar with the activity, a five minutes training run was conducted. Participants were instructed that they would not get paid for the bags produced in the training run. After the training run, the number of produced bags was recorded. These bags were collected and subsequently used for reforestation. This training run served several purposes. First, it assured homogeneous quality of produced bags. Second, the data are useful to calibrate baseline performance to reduce noise in the estimates. This is particularly helpful to identify – among other things – gender effects after treatment, because we can rule out that the performance was different before treatment due to reasons not related to the monetary incentives. After the training run, the facilitator provided the enumerator with eight closed envelopes which contained information on the payment participants would receive for each bag they produce (see 1.2.9). This procedure was chosen to make the treatment and potential pay differences more salient to participants. Then, enumerators showed all eight envelopes and, depending on the treatment, they told the participants that

- …in all eight envelopes, there will be a 600 Vietnamese Dong (VND) rate per bag (group 1)

- …in all eight envelopes, there will be a 400 VND rate per bag (group 2)
- …in four envelopes, there will be a 400 VND rate per bag and in four envelopes there will be a 600 VND rate per bag (groups 3 and 4)

Group members were asked to draw one of these closed envelopes and read about their payment. Illiterate people were helped by enumerators in private. To prevent communication between the groups, groups were placed at least four meters away from each other. Furthermore, participants were instructed not to talk with participants from other groups, which could be prevented succesfully. Communication within groups was allowed (85% of participants reported in the survey that they talked to others. This made the task more realistic and allowed inidividuals to talk about their payments, further increasing salience of the treatments (32% of participants reported in the survey that they talked to others about their pay rates; see 1.5 for summary by treatment).

Enumerators responsible for a group conducted the counting of prepared soil bags during and at the end of the task. Each group member who had produced 20 soil bags and placed them in her box was asked to get up, showed the box with the bags to the group's enumerator, unloaded the bags from her box, and went back to her seat to continue soil bag production. This way, the enumerator counted the bags produced by each subject ideally in blocks of 20 during the task. Yet, the enumerator did count each bag in the boxes as participants occasionally put fewer or more than 20 bags into their boxes.

After the 60 minutes of the task, subjects were instructed to stop working, and the enumerator counted the remaining number of bags that were placed in the individual box. The total number of bags produced was written down by the enumerator on a group's results sheet that was collected by the team of principal researchers.

After the task, participants remained seated in their groups and enumerators conducted the survey (see 1.3) by interviewing the participants, i.e. reading out the questions and marking the answers. At no point were names of the participants recorded.

For payments, each group's result sheet was handed over to an assistant not involved in the supervision of the task. This assistant calculated of the money earned (i.e., number of bags produced times pay rate plus the show-up fee). The assistant then prepared an envelope with an ID number that contained the cash payment for each subject. These envelopes were handed over to another assistant who was not involved in the group supervision to distribute them among participants. This way, the payment was double blind.

After the surveys were conducted and the payments were distributed, a debriefing by the research staff was offered. Besides the general topic of forestry in the area, the purpose of the

experiment, i.e., studying the effects of inequitable payments on the effectiveness of conservation policies, was revealed, explained, and discussed. To ensure the important design feature of tangible environmental benefits in our experiment, all soil bags were handed over to a local forest management organization which used them for reforestation activities in the region. This was known to all participants. Across all 14 study villages, 38445 soil bags were produced in our experiments, including those from training runs.

### 1.2.8 Consent form

The consent form had the following wording:

We want to conduct an activity with you as part of our research on local forestry. We offer you to take part in a task that can help to protect forests and you will be paid for it. Many different things are needed to protect forests. One important activity is to plant new trees which must be prepared. The work we will pay you for consists of the packing of soil bags for reforestation/conservation. Please note that you cannot take part in the study if you are below 18 years of age.

In total, our activity will last approximately 2.5 hours. By showing up here, you have already earned 30,000 Dong, but you can earn more if you participate in the activity. You will be paid per soil bag packed. Your participation is fully voluntary. You cannot lose any money. We record no personal identifying information like names, phone numbers etc. No information we collect in the activity can be linked to your personal identity. All data will be used only for research purposes and in anonymous form.

After the activity, we would like to ask you some questions in a questionnaire. This will also be anonymous, and you can skip any questions that you would not like to answer. After the activity, you will have the opportunity to talk about it with us. If you agree to participate, please raise your hand now, and sign the sheet that is prepared for you. If you would not like to participate, you can leave the activity now and keep the 30,000 Dong you have earned already. If you have further questions, please feel free to ask now.

Note that you can also leave the activity at any time after having agreed to participate. In that case, you can only keep the 30,000 Dong you have earned by now.

I have read (or it has been read out for me) and understood the content of the study, and I give my consent to participate.

### 1.2.9 Payment information sheets

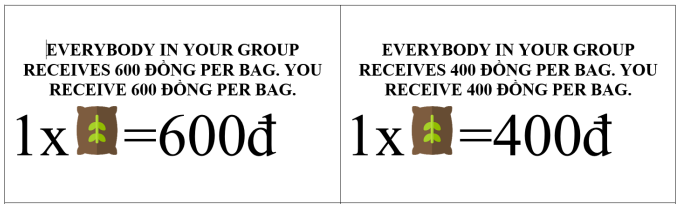


**Figure S2:** Information sheet about own payment and payment of others in a group with equal pay rates, handed to participants in closed envelopes (left: sheet for high pay rate; right: sheet for low pay rate).

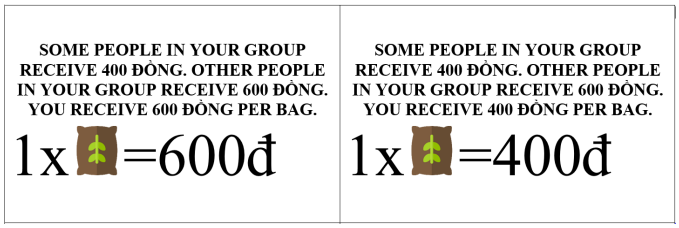


**Figure S3:** Information sheet about own payment and payment of others in a group with unequal pay rates, handed to participants in closed envelopes (left: sheet for high pay rate; right: sheet for low pay rate).

### 1.2.10 Instructions

Welcome everybody! We are a group of researchers, consisting of two researchers from Europe, and seven assistants. We will introduce them to you.

*- Introduce staff with names - Read out consent form - Show consent form and collect signatures - Distribute randomized envelopes with IDs to participants*

We will explain the task now for which we will pay you. Please raise your hand at any time if you have a question or are unsure whether you understood what we said. The task will be the preparation of soil bags. Empty bags need to be filled with soil such that seedlings can be put into them later and they can be planted. Thus, the bags will help local forest officers to rebuild forests. This makes it important that the bags are of good quality. We will show you how to prepare a soil bag. Your payment will depend on how many bags of good quality you produce.

*- Demonstrate soil bag preparation and storage in box*

Each one of you will work on this task. Next to your seat, you can see a cardboard box in which you will place the prepared bags.

Each of you packs the soil bags into the small boxes. Pack 20 soil bags in the small box. Then you go to the assistants and show their box with the 20 soil bags and your ID number to the

assistant. The assistant will note the number. Then you pack the 20 soil bags from the small box into the big box and show the empty small box to the assistants. You go back to your place and pack the next soil bags. You will have 60 minutes to work on the preparation of such bags to earn money. After the 60 minutes, we will collect all of your bags, count them, and pay you with the amount you earned for the bags you prepared. We will carefully count the bags which you produce, so you do not have to take care of this. The bags will be given to the local forest officials. Importantly, nobody, neither we nor them, will know who produced how many bags, because we do not record your names. It is fully up to you how much of the task time you want to work to earn money. We will only count the bags you produced. Of course, you can take breaks if you want to and have drinks.

You will be split into four smaller groups for the task. The groups will be seated some distance away from each other, so you can chat with people in your group while working, but not with people who are in other groups. Remember that although you work in a group, each of you will be paid individually and only for the bags produced by yourself. It does not matter for your payment how many bags the others in your group prepare. We will show you a sheet of paper where you can see which IDs are in which group. After you gathered into the four groups by showing your ID to us, please follow the assistant to the working place with your number. When you are seated in your group, you will get information on the payment which you receive for each bag you prepare.

*- Split participants into groups*

We will now conduct a short training phase of soil bag packing for 5 minutes. You are not getting paid for the bags you prepare within these 5 minutes. We just want everybody to get familiar with the task and want you to have the opportunity to learn how to prepare bags of good quality.

*- Conduct training run*

We will now distribute envelopes in which you find information on the payment you will receive from us for each bag you produce.

*- Show closed envelopes to group - Tell about payment rate distribution in the group*

The payment information in your envelopes applies to you, independent of what other people have in their envelopes.

*- Distribute randomized payment rate envelopes within groups by letting them draw - Remind participants that they can talk within the group, not between groups - Remind participants of water and snacks and that they can take breaks - Run task for 60 minutes*

While we calculate your payments according to the number of bags that have been packed, we

would ask you to participate in a brief survey where you answer some questions. The assistants will come to you individually. Let us know when you do not want to respond to a question.

## 1.3 Survey

For each participant in the experiment, village, date and ID were recorded and the following questions were asked. Question 23 presents vignettes adapted from (28) which were further explained with visual aids (Figure S4).

1. Please indicate your gender
2. How old are you?
3. Which ethnicity do you belong to?
4. What is a normal month's income for you?
5. What is your major source of income?
6. How many years of formal education did you receive?
7. Please indicate which of the following items can be found in your household
    - Motorcycle
    - Cell phone
    - TV
    - Fridge
    - Gas stove
8. How many children (15 years or younger) are living in your household?
9. Have you prepared seedling bags on another occasion before?
10. How difficult/exhausting was the preparation of seedling bags for you?
    - Very difficult
    - Somewhat difficult
    - Undecided
    - Not so difficult
    - Not difficult at all
11. How did you like the snacks and drinks we provided?
    - Very much
    - Somewhat
    - Undecided
    - Not so much
    - Not at all

12. What do you think, how much time of the 60 minutes did you spend *not* filling bags, but doing something else?
13. Did you chat with the other people while you were working?
14. If you chatted with the other people while working, did you talk about how much you and others earn?
15. Did you know the pay rate of all other people in your group?
16. How fair do you think your own payment rate was?

    - Very fair
    - Somewhat fair
    - Undecided
    - Not so fair
    - Not fair at all

17. Why and what do you think about it?
18. How worried are you about the future of the forest?

    - Not worried at all
    - Not so worried
    - Undecided
    - Somewhat worried
    - Very worried

19. How important is the forest for the future of you and your family?

    - Not important at all
    - Not so important
    - Undecided
    - Somewhat important
    - Very important

20. Do you take part in patrolling, monitoring, firefighting or any other activities for the protection of the forest?
21. Are there people in your village who are logging trees from the forest?
22. Are there people in your village who go hunting in the forest?
23. Imagine the following four scenarios:

    - Person A and person B are identical in terms of physical and mental abilities. They become lost in an uninhabited forest where the only food is mango. They can collect as many mangos as they want by picking them from the trees and throwing them into a pile. In this way, person A picks 12 mangos and person B picks 8 mangos per

day. Person A takes from the pile the 12 mangos he picked leaving person B with the 8 which B picked. This is

  - Fair *(Note for reader: this response is consistent with the accountability principle)*
  - Unfair

- Person A and Person B become lost in an uninhabited forest where the only food is mangos. They can collect as many mangos as they want by picking them from the trees and throwing them into a pile. A and B are identical in terms of physical and mental abilities except that A was born with one hand and B with two. Together they pick a total of 20 mangos per day, but because of his condition A picks fewer mangos per day than B. B takes 12 mangos from the pile leaving 8 for A. This is
  - Fair
  - Unfair *(Note for reader: this response is consistent with the accountability principle)*
- Person A and person B become lost in an uninhabited forest where the only food is mangos. They can collect as many mangos as they want by picking them from the trees and throwing them into a pile. Person A and person B are identical in terms of physical and mental abilities except that person A was born with one hand and person B with two. Together they pick a total of 20 mangos per day, but because of his condition person A picks fewer mangos per day than person B. B takes 10 mangos from the pile leaving 10 for A. This is
  - Fair *(Note for reader: this response is consistent with the accountability principle)*
  - Unfair
- Person A and person B become lost in an uninhabited forest where the only food is mangos. They can collect as many mangos as they want by picking them from the trees and throwing them into a pile. The two men are identical in terms of physical and mental abilities except that person A takes 5 hours to pick the mangos while B takes 4 hours because person A chooses to work at a slower pace. Nevertheless, they pick 9 mangos each for the total of 18. A takes 10 mangos from the pile leaving 8 for B. This is
  - Fair
  - Unfair *(Note for reader: this response is consistent with the accountability principle)*

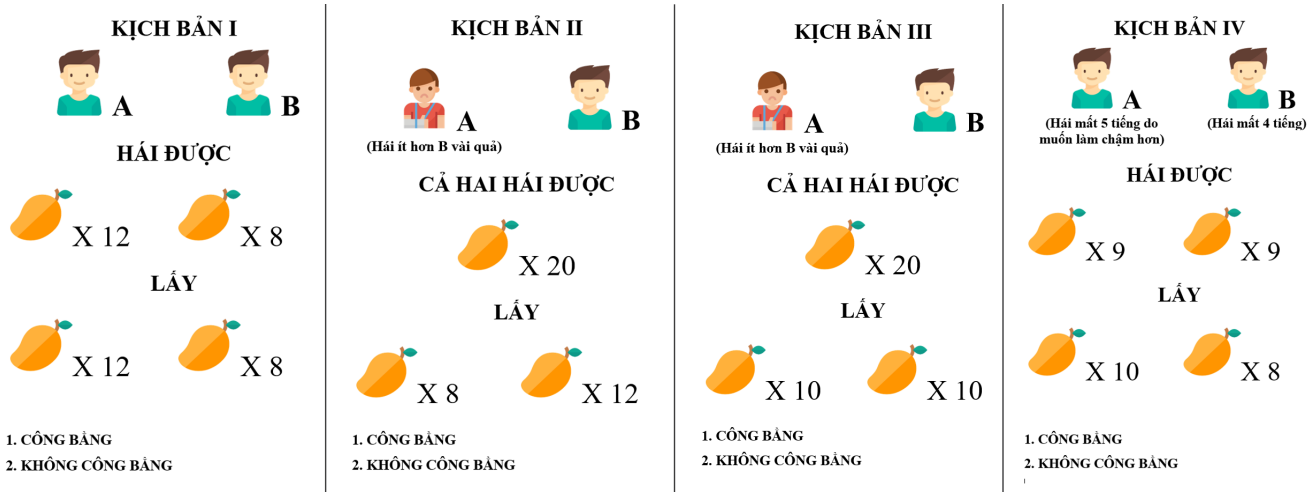


**Figure S4:** Visual aids for vignettes 1-4 of question 23 in the survey (Vietnamese).

## 1.4 Pre-registration and ethical approval

All major hypotheses on conservation effort, model specification, the set of control variables, a power analysis and the outlier adjustment method were pre-registered (2019-03-22) before data collection at the pilot site started (2019-03-24, see Table S1). Ethical approval was obtained.

Link to pre-registration document at AsPredicted.org:

*http://aspredicted.org/blind.php?x=k4ij32*

Note that we deviated from the pre-registration in three small aspects. First, data was collected in only 14 instead of 15 randomly selected villages (holding constant the pre-specified total sample size of subjects), because one village served as pilot for testing and training the procedures in the field. We employed no treatment manipulations in the pilot village (see 1.2.1). Second, prior to analysis, we chose to model between-village variation with random effects (fitted via Restricted Maximum Likelihood approximation) instead of fixed effects. This is preferable when the predictor (here: treatment) is uncorrelated with the cluster (here: village), which we achieved due to randomization at the village level. Further, village-specific intercepts in our data can be viewed as a random sample from a larger population of village-specific intercepts (due to the random village selection procedure, see 1.2.1). Random effects use this information by modelling village-level intercepts as draws from a Normal distribution (shrinkage), improving efficiency. Third, instead of employing the variables *Forest worrying* and *Forest importance* (both measured in the survey, see 1.3)), as separate covariates variables in the respective models, we use the sum of their standardized scores as a measure of *Forest concern* (see Models 3, 6, 7, 8 in Tables S11 to S14 in 2.1.3). This was decided post hoc due to their high collinearity ($r = 0.42$) and the small variance in participants' responses (see Table S4). It does not affect

estimated treatment effects.

Link to ethical approval at German Association for Experimental Economic Research e.V.:

*https://gfew.de/ethik/FyVXvL4E*

## 1.5 Descriptive statistics for villages and participants

Table S3 shows village information. Tables S4 and S5 show descriptive statistics for all survey variables used in the analyses, training run performance and payoff, both aggregated and split by treatment, respectively. As expected under randomization, all basic socio-demographic variables were balanced across treatments as indicated by p-values from ANOVA (continuous variables) and Chi-square tests (categorical variables) for gender ($p = 0.7$), age ($p = 0.47$), education ($p = 0.77$), children ($p = 0.22$), income ($p = 0.75$) and posession of household items motorcycle ($p = 0.79$), cell phone ($p = 0.43$), TV ($p = 0.59$), fridge ($p = 0.77$) and gas stove ($p = 0.88$).

**Table S3:** Village information for the 14 experimental sites and one pilot site (first row). Village information was collected in January and February 2019 from communal administrative records (PFES: Payments for Forest Ecosystem Services).

| Village | Commune | No. of households | Population | Ethnic majority | Avg. daily income (USD) | Poverty rate (%) | Forest area under PFES (ha) |
|---|---|---|---|---|---|---|---|
| 8 | Huong Hoa | 206 | 839 | Kinh | 1.6 | 1.4 | 11.2 |
| Cha Mang | Thuong Lo | 82 | 324 | Katu | 1.2 | 4.8 | 694.7 |
| 6 | Thuong Long | 76 | 307 | Katu | 1.2 | 46.7 | 62.8 |
| 8 | Thuong Long | 104 | 480 | Katu | 1.2 | 27.0 | 66.3 |
| La Van | Thuong Nhat | 86 | 354 | Katu | 1.3 | 5.9 | 91.3 |
| Ta Lu | Thuong Nhat | 67 | 270 | Katu | 1.3 | 4.8 | 82.0 |
| Ria Ho | Thuong Lo | 78 | 336 | Katu | 1.2 | 3.8 | 769.0 |
| Ta Rung | Huong Son | 108 | 447 | Katu | 1.3 | 3.7 | 63.3 |
| Bha Bha | Huong Son | 96 | 388 | Katu | 1.3 | 5.2 | 99.5 |
| 1 | Thuong Quang | 94 | 366 | Katu | 1.4 | 6.3 | 117.5 |
| 5 | Thuong Quang | 88 | 338 | Kinh | 1.4 | 5.6 | 132.2 |
| 6 | Thuong Quang | 71 | 273 | Kinh | 1.4 | 1.4 | 157.0 |
| Ta Rinh | Thuong Nhat | 108 | 460 | Katu | 1.3 | 5.6 | 161.4 |
| A2 | Huong Son | 75 | 307 | Katu | 1.3 | 4.0 | 76.5 |
| Pa Noong | Huong Son | 83 | 364 | Katu | 1.3 | 2.4 | 80.0 |

**Table S4:** Summary of descriptive statistics for all survey variables used in the analyses, training run performance and payoff. Instead of the score 'Household wealth' with mean = 0 and SD = 1 from a principal component analysis (PCA), household items that went into PCA are shown one by one. The two separate variables 'Forest worrying' and 'Forest importance' are shown instead of the sum of their standardized scores 'Forest concern'.

| Variable | Result | Survey question |
|---|---|---|
| N | 443 | - |
| Female (N, %) | 265 (59.8) | 1 |
| Age (mean, SD) | 38.50 (12.96) | 2 |
| Education (years) (mean, SD) | 6.45 (4.45) | 6 |
| Children (mean, SD) | 1.20 (0.95) | 8 |
| Income (USD per month) (mean, SD) | 82.85 (80.88) | 4 |
| Household item: Motorcycle (N, %) | 372 (84.0) | 7 |
| Household item: Cell phone (N, %) | 381 (86.0) | 7 |
| Household item: TV (N, %) | 390 (88.0) | 7 |
| Household item: Fridge (N, %) | 133 (30.0) | 7 |
| Household item: Gas stove (N, %) | 171 (38.6) | 7 |
| **Forest worrying (N, %)** | | |
| Not so much | 7 (1.6) | 18 |
| Undecided | 5 (1.1) | 18 |
| Somewhat | 78 (17.6) | 18 |
| Very much | 353 (79.7) | 18 |
| **Forest importance (N, %)** | | |
| Not so much | 1 (0.2) | 19 |
| Undecided | 4 (0.9) | 19 |
| Somewhat | 21 (4.7) | 19 |
| Very much | 417 (94.1) | 19 |
| **Likeability of snacks (N, %)** | | |
| Not at all | 7 (1.6) | 11 |
| Not so much | 11 (2.5) | 11 |
| Undecided | 9 (2.1) | 11 |
| Somewhat | 110 (25.1) | 11 |
| Very much | 302 (68.8) | 11 |
| **Fairness own pay rate (N, %)** | | |
| Not at all | 15 (3.4) | 16 |
| Not so much | 74 (16.7) | 16 |
| Undecided | 25 (5.7) | 16 |
| Somewhat | 220 (49.8) | 16 |
| Very much | 108 (24.4) | 16 |
| Reported time not working (min) (mean, SD) | 3.06 (5.16) | 12 |
| Prior experience task (N, %) | 58 (13.2) | 9 |
| Endorsement of accountability principle (0-4) (mean, SD) | 3.19 (0.73) | 23 |
| Participation in forest protection (N, %) | 286 (64.6) | 20 |
| Talked to others in group about pay rate (N, %) | 140 (31.7) | 14 |
| Training run (mean, SD) | 6.42 (2.24) | - |
| Payoff (USD) (mean, SD) | 3.01 (0.52) | - |

**Table S5:** Summary of descriptive statistics split by treatment.

| | Equal, High pay rate | Equal, Low pay rate | Unequal, High pay rate | Unequal, Low pay rate |
|---|---|---|---|---|
| N | 112 | 110 | 109 | 112 |
| Female (N) | 62 | 69 | 67 | 67 |
| Age (mean, SD) | 38.6 (13) | 40 (13.5) | 38.1 (13.2) | 37.3 (12.2) |
| Education (years) (mean, SD) | 6.3 (4.5) | 6.2 (4.2) | 6.8 (4.6) | 6.5 (4.5) |
| Children (mean, SD) | 1.1 (0.9) | 1.2 (1) | 1.2 (1) | 1.3 (0.9) |
| Income (USD per month) (mean, SD) | 90.1 (93) | 79.9 (73) | 81.6 (78.9) | 79.8 (77.9) |
| Household item: Motorcycle (N) | 96 | 90 | 90 | 96 |
| Household item: Cell phone (N) | 101 | 92 | 91 | 97 |
| Household item: TV (N) | 100 | 93 | 96 | 101 |
| Household item: Fridge (N) | 32 | 31 | 32 | 38 |
| Household item: Gas stove (N) | 46 | 43 | 39 | 43 |
| **Forest worrying (N)** | | | | |
| Forest worrying - Not so much | 1 | 1 | 2 | 3 |
| Forest worrying - Undecided | 3 | 0 | 1 | 1 |
| Forest worrying - Somewhat | 22 | 15 | 19 | 22 |
| Forest worrying - Very much | 86 | 94 | 87 | 86 |
| **Forest importance (N)** | | | | |
| Forest importance - Not so much | 0 | 0 | 0 | 1 |
| Forest importance - Undecided | 3 | 1 | 0 | 0 |
| Forest importance - Somewhat | 3 | 8 | 4 | 6 |
| Forest importance - Very much | 106 | 101 | 105 | 105 |
| **Likeability of snacks (N)** | | | | |
| Likeability of snacks - Not at all | 2 | 0 | 2 | 3 |
| Likeability of snacks - Not so much | 1 | 3 | 3 | 4 |
| Likeability of snacks - Undecided | 2 | 1 | 3 | 3 |
| Likeability of snacks - Somewhat | 31 | 18 | 28 | 33 |
| Likeability of snacks - Very much | 76 | 86 | 71 | 69 |
| **Fairness own pay rate (N)** | | | | |
| Fairness own pay rate - Not at all | 1 | 5 | 3 | 6 |
| Fairness own pay rate - Not so much | 6 | 26 | 11 | 31 |
| Fairness own pay rate - Undecided | 8 | 4 | 6 | 7 |
| Fairness own pay rate - Somewhat | 50 | 61 | 53 | 56 |
| Fairness own pay rate - Very much | 47 | 13 | 36 | 12 |
| Reported time not working (min) (mean, SD) | 3.2 (3.6) | 3.6 (7.9) | 2.3 (2.8) | 3.1 (4.8) |
| Prior experience task (N) | 16 | 11 | 14 | 17 |
| Endorsement of accountability principle (0-4) (mean, SD) | 3.2 (0.7) | 3 (0.7) | 3.2 (0.8) | 3.3 (0.7) |
| Participation in forest protection (N) | 71 | 63 | 83 | 69 |
| Talked to others in group about pay rate (N) | 35 | 42 | 36 | 27 |
| Training run (mean, SD) | 6.5 (2.2) | 6.8 (2.3) | 6.1 (2.3) | 6.3 (2.1) |
| Payoff (USD) (mean, SD) | 3.3 (0.5) | 2.8 (0.4) | 3.3 (0.5) | 2.7 (0.3) |

## 1.6 Statistical outliers

Real-effort experiments are prone to high variance in effort compared to stated-effort experiments (1), which makes extreme outcome values more likely. This is because, unlike in stated-effort experiments, the costs of physical effort vary widely across participants with different abilities, health, curiosity or enjoyment of the activity, which can compromise inference. Based on this and previous data with the task (6), we defined what counts as an outlier in the pre-registration (see 1.4). We considered outliers those observations of conservation effort that fell outside *Q1 - 1.5* x *IQR* and *Q3 + 1.5* x *IQR* of a treatment's distribution, *Q1(Q3)* referring to the upper limit of the first (third) quartile and *IQR* to inter-quartile range (29). As planned, we ran all models without outlier adjustment (*including* outlier observations) and with outlier adjustment (*excluding* outlier observations). Extreme values can have multiple reasons, and it is important to look for further justifications for inlcusion or exclusion in effect estimation

(29). Table S6 provides further information on all subjects who were statistically classified as outliers: treatment assignment, effort difference from treatment median, comments taken by enumerators in the field and, as a measure of influence of the observation on regression fit, Cook's distance (30). Cook's distance for mixed-effect models was calculated with the package *influence.ME* (31). Inspecting enumerators' field notes and Cook's distances with relation to the conventional cutoff for highly influential observations (Cook's distance $> 4/N = 4/443 = 0.009$; (31)) supports the choice to base major conclusions about treatment effects on results *with* outlier adjustment (results in main text). Conclusions do not qualitatively change when including all raw data (see 2.1.3), but note that the effect size of the unequal treatment in low-paid subjects is smaller without outlier adjustment and less precisely estimated when we control for group ID. The smaller effect size is mainly due to three strong outliers in the equal groups with low pay rate (see uppermost panel in Figure 3 in main text) who entirely stopped working for reasons that are most likely not related to treatment (see Table S6).

**Table S6:** Description of outlier observations of conservation effort (No. of bags).

| ID | Treatment | Diff. from median | Cook's dist. | Comments on participant by enumerators |
|---|---|---|---|---|
| 32 | Unequal, Low pay rate | 60.5 | 0.011 (> 0.009) | produced soil bags quickly |
| 34 | Unequal, High pay rate | 60.0 | 0.012 (> 0.009) | (no comment) |
| 40 | Unequal, Low pay rate | 60.5 | 0.011 (> 0.009) | produced soil bags quickly |
| 116 | Equal, High pay rate | 57.0 | 0.014 (> 0.009) | did the first small box with 20 soil bags in 10 mins already |
| 246 | Equal, High pay rate | -50.0 | 0.017 (> 0.009) | was the last one due to hand pain |
| 249 | Equal, Low pay rate | -78.0 | 0.043 (> 0.009) | did not want to work, stopped working after 12 minutes |
| 308 | Equal, High pay rate | 52.0 | 0.015 (> 0.009) | is fastest worker with 20 soil bags in 8 min |
| 313 | Equal, Low pay rate | -63.0 | 0.027 (> 0.009) | stopped working because his back got tired |
| 350 | Equal, Low pay rate | -63.0 | 0.027 (> 0.009) | stopped working after 16 min, had helped to prepare the soil for money, thought was enough with show-up fee |
| 365 | Unequal, Low pay rate | -59.5 | 0.023 (> 0.009) | using phone after 3 mins, was first to eat snacks after 27 min and stopped doing soil bags |

# 2 Supplementary Results

## 2.1 Treatment effects on effort

### 2.1.1 Summary of effort by treatment

**Table S7:** Summary statistics for conservation effort (No. of bags) by treatment. No outlier adjustment (N = 443).

| Treatment | Mean | SD | Median | Q1 | Q3 | Min | Max | N |
|---|---|---|---|---|---|---|---|---|
| Equal, High pay rate | 77.1 | 19.5 | 78.0 | 63.8 | 86.5 | 28 | 135 | 112 |
| Equal, Low pay rate | 85.6 | 20.6 | 83.0 | 77.0 | 100.0 | 5 | 123 | 110 |
| Unequal, High pay rate | 78.4 | 19.2 | 79.0 | 64.0 | 91.0 | 40 | 139 | 109 |
| Unequal, Low pay rate | 80.5 | 20.0 | 79.5 | 68.8 | 93.2 | 20 | 140 | 112 |
| Unequal (aggregated) | 79.4 | 19.6 | 79.0 | 67.0 | 93.0 | 20 | 140 | 221 |
| Equal (aggregated) | 81.3 | 20.4 | 80.0 | 68.0 | 94.0 | 5 | 135 | 222 |
| All | 80.4 | 20.0 | 80.0 | 67.5 | 93.0 | 5 | 140 | 443 |

**Table S8:** Summary statistics for conservation effort (No. of bags) by treatment. With outlier adjustment (N = 433).

| Treatment | Mean | SD | Median | Q1 | Q3 | Min | Max | N |
|---|---|---|---|---|---|---|---|---|
| Equal, High pay rate | 76.5 | 17.6 | 78 | 64.0 | 85.0 | 40 | 120 | 109 |
| Equal, Low pay rate | 87.6 | 17.1 | 83 | 77.5 | 100.0 | 55 | 123 | 107 |
| Unequal, High pay rate | 77.9 | 18.4 | 79 | 63.8 | 90.2 | 40 | 123 | 108 |
| Unequal, Low pay rate | 79.9 | 17.6 | 79 | 69.0 | 92.0 | 36 | 126 | 109 |
| Unequal (aggregated) | 78.9 | 18.0 | 79 | 67.0 | 91.0 | 36 | 126 | 217 |
| Equal (aggregated) | 82.0 | 18.2 | 80 | 69.5 | 94.0 | 40 | 123 | 216 |
| All | 80.4 | 18.1 | 80 | 68.0 | 93.0 | 36 | 126 | 433 |

### 2.1.2 Paired treatment comparisons

**Table S9:** Results from Welch's t-tests, randomization tests (5,000 permutations), and Wilcoxon rank sum tests for all treatment comparisons. No outlier adjustment (N = 443).

| Comparison | t-test | Randomization test | Wilcoxon test |
|---|---|---|---|
| Equal, High pay rate vs Equal, Low pay rate | t = -3.14, p = 0.0019 | p = 0.0024 | p = 2e-04 |
| Equal, High pay rate vs Unequal, Low pay rate | t = -1.27, p = 0.2050 | p = 0.2110 | p = 0.1930 |
| Equal, High pay rate vs Unequal, High pay rate | t = -0.5, p = 0.6161 | p = 0.6000 | p = 0.5440 |
| Equal, Low pay rate vs Unequal, Low pay rate | t = 1.88, p = 0.0617 | p = 0.0592 | p = 0.0119 |
| Equal, Low pay rate vs Unequal, High pay rate | t = 2.66, p = 0.0084 | p = 0.0108 | p = 0.0019 |
| Unequal, Low pay rate vs Unequal, High pay rate | t = 0.78, p = 0.4390 | p = 0.4380 | p = 0.4929 |
| Equal (aggregated) vs Unequal (aggregated) | t = 0.97, p = 0.3302 | p = 0.3412 | p = 0.1967 |

**Table S10:** Results from Welch's t-tests, randomization tests (5,000 permutations), and Wilcoxon rank sum tests for all treatment comparisons. With outlier adjustment (N = 433).

| Comparison | t-test | Randomization test | Wilcoxon test |
|---|---|---|---|
| Equal, High pay rate vs Equal, Low pay rate | t = -4.67, p = 0.0000 | p = 0.0000 | p = 0.0000 |
| Equal, High pay rate vs Unequal, Low pay rate | t = -1.42, p = 0.1581 | p = 0.1636 | p = 0.1762 |
| Equal, High pay rate vs Unequal, High pay rate | t = -0.54, p = 0.5921 | p = 0.5832 | p = 0.5358 |
| Equal, Low pay rate vs Unequal, Low pay rate | t = 3.24, p = 0.0014 | p = 0.0012 | p = 0.0019 |
| Equal, Low pay rate vs Unequal, High pay rate | t = 4.01, p = 1e-04 | p = 0.0000 | p = 2e-04 |
| Unequal, Low pay rate vs Unequal, High pay rate | t = 0.85, p = 0.3988 | p = 0.4076 | p = 0.4810 |
| Equal (aggregated) vs Unequal (aggregated) | t = 1.79, p = 0.0744 | p = 0.0740 | p = 0.0939 |

### 2.1.3 Full modelling results and alternative specifications

Full regression results are reported for the data without outlier adjustment (Table S11), with outlier adjustment (Table S12, as in main text), and also with outlier adjustment plus an additional random effect for group ID (Table S13) and with clustered standard errors (using the *plm* package (32)) on the group level (Table S14). This is because in each village there are two groups with identical (unequal) treatment assignment. For details on outlier definition and handling, see 1.6. In addition to models with the set of predictors as shown in the main text (Models 1, 3, 4, 6 and 8 in Tables S11 to S14), specifications with a smaller set of socio-demographic controls only are shown (Models 2 and 5 in Tables S11 to S14), as well as specifications which estimate the aggregate effect of the unequal treatment across different pay rates (i.e., ignoring the distinction between advantageous and disadvantageous inequality; Model 7 in Tables S11 to S14). For a statistical overview of control variables see 1.5.

**Table S11:** Results from linear mixed-effect models with village as random effect. Dependent and independent variables (except binary variables) are standardized to have mean = 0 and standard deviation = 1. No outlier adjustment.

| | Low pay rate | | | High pay rate | | | All | |
|---|---|---|---|---|---|---|---|---|
| Variable | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
| Unequal | -0.249 (0.125)** | -0.275 (0.115)** | -0.191 (0.095)** | 0.062 (0.118) | 0.035 (0.112) | 0.135 (0.09) | -0.018 (0.066) | -0.164 (0.093)* |
| Female | | 0.616 (0.131)*** | 0.482 (0.106)*** | | 0.594 (0.135)*** | 0.374 (0.107)*** | 0.407 (0.076)*** | 0.398 (0.076)*** |
| Age | | -0.061 (0.07) | -0.102 (0.057)* | | -0.086 (0.071) | -0.097 (0.057)* | -0.097 (0.04)** | -0.102 (0.04)** |
| Wealth | | 0.198 (0.071)*** | 0.122 (0.057)** | | 0.114 (0.07) | 0.048 (0.055) | 0.081 (0.039)** | 0.087 (0.039)** |
| Education | | -0.038 (0.072) | 0.015 (0.057) | | 0.059 (0.073) | 0.075 (0.058) | 0.041 (0.041) | 0.036 (0.041) |
| Children | | 0.144 (0.061)** | 0.089 (0.049)* | | -0.026 (0.063) | 0.004 (0.05) | 0.038 (0.035) | 0.039 (0.035) |
| Income | | -0.101 (0.067) | -0.115 (0.053)** | | 0.145 (0.065)** | 0.086 (0.052) | -0.015 (0.037) | -0.015 (0.037) |
| Forest concern | | | 0.007 (0.05) | | | 0.027 (0.048) | 0.034 (0.034) | 0.03 (0.034) |
| Training run | | | 0.629 (0.054)*** | | | 0.611 (0.054)*** | 0.621 (0.038)*** | 0.619 (0.038)*** |
| Prior experience task | | | 0.042 (0.15) | | | 0.16 (0.142) | 0.102 (0.103) | 0.118 (0.103) |
| Likeability of snacks | | | -0.047 (0.049) | | | 0.118 (0.048)** | 0.036 (0.034) | 0.031 (0.034) |
| High pay rate | | | | | | | -0.174 (0.065)*** | -0.319 (0.092)*** |
| Unequal x high pay rate | | | | | | | | 0.29 (0.13)** |
| N | 222 | 221 | 217 | 221 | 218 | 216 | 433 | 433 |
| Marginal R^2 | 0.02 | 0.16 | 0.51 | 0 | 0.11 | 0.48 | 0.49 | 0.49 |
| Conditional R^2 | 0.15 | 0.37 | 0.56 | 0.25 | 0.41 | 0.57 | 0.55 | 0.55 |
| Log likelihood | -971.9 | -947.3 | -865.1 | -946.3 | -920.5 | -852.5 | -1734.5 | -1730.2 |

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

**Table S12:** Results from linear mixed-effect models with village as random effect. Dependent and independent variables (except binary variables) are standardized to have mean = 0 and standard deviation = 1. With outlier adjustment.

| | Low pay rate | | | High pay rate | | | All | |
|---|---|---|---|---|---|---|---|---|
| Variable | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
| Unequal | -0.423 (0.119)*** | -0.428 (0.111)*** | -0.351 (0.092)*** | 0.062 (0.118) | 0.029 (0.114) | 0.119 (0.094) | -0.082 (0.067) | -0.291 (0.094)*** |
| Female | | 0.564 (0.127)*** | 0.51 (0.103)*** | | 0.543 (0.137)*** | 0.365 (0.112)*** | 0.411 (0.077)*** | 0.398 (0.076)*** |
| Age | | -0.07 (0.068) | -0.102 (0.055)* | | -0.108 (0.072) | -0.123 (0.06)** | -0.114 (0.04)*** | -0.119 (0.04)*** |
| Wealth | | 0.189 (0.068)*** | 0.125 (0.055)** | | 0.077 (0.071) | 0.048 (0.058) | 0.084 (0.04)** | 0.091 (0.04)** |
| Education | | -0.028 (0.069) | 0.037 (0.056) | | 0.081 (0.073) | 0.085 (0.06) | 0.061 (0.041) | 0.054 (0.041) |
| Children | | 0.098 (0.059)* | 0.071 (0.047) | | -0.066 (0.063) | -0.022 (0.052) | 0.018 (0.036) | 0.019 (0.035) |
| Income | | 0.079 (0.064) | 0.012 (0.051) | | 0.111 (0.066)* | 0.061 (0.055) | 0.022 (0.038) | 0.023 (0.037) |
| Forest concern | | | -0.071 (0.048) | | | 0.023 (0.05) | -0.008 (0.035) | -0.013 (0.034) |
| Training run | | | 0.585 (0.054)*** | | | 0.574 (0.057)*** | 0.588 (0.039)*** | 0.583 (0.038)*** |
| Prior experience task | | | 0.087 (0.146) | | | 0.142 (0.148) | 0.129 (0.105) | 0.157 (0.104) |
| Likeability of snacks | | | -0.134 (0.048)*** | | | 0.119 (0.051)** | 0.004 (0.034) | -0.004 (0.034) |
| High pay rate | | | | | | | -0.243 (0.066)*** | -0.45 (0.093)*** |
| Unequal x high pay rate | | | | | | | | 0.411 (0.131)*** |
| N | 216 | 215 | 212 | 217 | 214 | 212 | 424 | 424 |
| Marginal R^2 | 0.04 | 0.16 | 0.5 | 0 | 0.09 | 0.43 | 0.46 | 0.47 |
| Conditional R^2 | 0.25 | 0.41 | 0.59 | 0.27 | 0.41 | 0.54 | 0.54 | 0.55 |
| Log likelihood | -905.8 | -884.2 | -811.7 | -912.2 | -889.1 | -829.4 | -1661.6 | -1654.9 |

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

**Table S13:** Results from linear mixed-effect models with village and group ID as random effects. Dependent and independent variables (except binary variables) are standardized to have mean = 0 and standard deviation = 1. With outlier adjustment.

| | Low pay rate | | | High pay rate | | | All | |
|---|---|---|---|---|---|---|---|---|
| Variable | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
| Unequal | -0.433 (0.164)** | -0.436 (0.159)** | -0.354 (0.128)** | 0.074 (0.186) | 0.041 (0.17) | 0.124 (0.126) | -0.084 (0.102) | -0.3 (0.127)** |
| Female | | 0.575 (0.12)*** | 0.513 (0.098)*** | | 0.424 (0.134)*** | 0.276 (0.112)** | 0.378 (0.073)*** | 0.372 (0.073)*** |
| Age | | -0.069 (0.065) | -0.103 (0.052)* | | -0.115 (0.069)* | -0.131 (0.059)** | -0.122 (0.038)*** | -0.125 (0.038)*** |
| Wealth | | 0.165 (0.066)** | 0.133 (0.053)** | | 0.023 (0.069) | 0.034 (0.057) | 0.082 (0.038)** | 0.086 (0.038)** |
| Education | | 0.01 (0.067) | 0.051 (0.054) | | 0.095 (0.07) | 0.078 (0.059) | 0.058 (0.039) | 0.053 (0.039) |
| Children | | 0.084 (0.059) | 0.052 (0.048) | | -0.048 (0.061) | -0.015 (0.051) | 0.017 (0.035) | 0.018 (0.035) |
| Income | | 0.081 (0.062) | -0.003 (0.05) | | 0.108 (0.065)* | 0.05 (0.055) | 0.021 (0.037) | 0.022 (0.037) |
| Forest concern | | | -0.085 (0.048)* | | | 0.048 (0.049) | 0.002 (0.033) | -0.002 (0.033) |
| Training run | | | 0.596 (0.055)*** | | | 0.577 (0.059)*** | 0.569 (0.041)*** | 0.567 (0.04)*** |
| Prior experience task | | | 0.121 (0.14) | | | 0.077 (0.15) | 0.135 (0.101) | 0.151 (0.101) |
| Likeability of snacks | | | -0.138 (0.048)*** | | | 0.115 (0.05)** | 0.007 (0.035) | 0.002 (0.035) |
| High pay rate | | | | | | | -0.151 (0.074)** | -0.464 (0.137)*** |
| Unequal x high pay rate | | | | | | | | 0.426 (0.162)** |
| N | 216 | 215 | 212 | 217 | 214 | 212 | 424 | 424 |
| Marginal R^2 | 0.05 | 0.16 | 0.51 | 0 | 0.07 | 0.42 | 0.43 | 0.45 |
| Conditional R^2 | 0.35 | 0.5 | 0.66 | 0.41 | 0.48 | 0.57 | 0.6 | 0.6 |
| Log likelihood | -901.9 | -879.5 | -807.2 | -904.1 | -883.8 | -827.2 | -1650.6 | -1645.3 |

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

**Table S14:** Results from linear mixed-effect models with group ID as random effect and standard errors clustered at the group level. Variables are not standardized. With outlier adjustment.

| | Low pay rate | | | High pay rate | | | All | |
|---|---|---|---|---|---|---|---|---|
| Variable | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 | Model 7 | Model 8 |
| Intercept | 87.757 (2.488)*** | 80.883 (5.994)*** | 65.047 (6.277)*** | 76.616 (3.31)*** | 75.707 (8.08)*** | 35.587 (9.251)*** | 50.774 (5.773)*** | 54.641 (5.486)*** |
| Unequal | -7.781 (3.481)** | -7.883 (3.627)** | -6.102 (2.528)** | 1.376 (4.135) | 0.868 (4.141) | 2.382 (2.713) | -1.431 (2.113) | -5.296 (2.534)** |
| Female | | 9.629 (1.969)*** | 8.716 (1.677)*** | | 6.106 (2.817)** | 3.685 (2.565) | 6.342 (1.514)*** | 6.226 (1.518)*** |
| Age | | -0.089 (0.079) | -0.138 (0.056)** | | -0.155 (0.103) | -0.182 (0.074)** | -0.171 (0.052)*** | -0.175 (0.05)*** |
| Wealth | | 4.214 (1.635)** | 3.788 (1.424)*** | | -0.247 (2.374) | 0.537 (1.785) | 2.221 (1.182)* | 2.318 (1.185)* |
| Education | | 0.098 (0.276) | 0.196 (0.224) | | 0.377 (0.3) | 0.252 (0.234) | 0.197 (0.151) | 0.181 (0.15) |
| Children | | 1.617 (0.87)* | 0.882 (0.703) | | -0.801 (1.086) | -0.191 (0.965) | 0.348 (0.604) | 0.368 (0.602) |
| Income | | 0.023 (0.018) | -0.001 (0.012) | | 0.024 (0.015) | 0.01 (0.01) | 0.005 (0.009) | 0.005 (0.009) |
| Forest concern | | | -0.968 (0.738) | | | 0.52 (0.605) | -0.059 (0.474) | -0.091 (0.467) |
| Training run | | | 5.058 (0.492)*** | | | 5.057 (0.427)*** | 4.895 (0.323)*** | 4.874 (0.322)*** |
| Prior experience task | | | 2.275 (1.697) | | | 0.548 (3.215) | 2.058 (1.863) | 2.282 (1.889) |
| Likeability of snacks | | | -3.037 (0.87)*** | | | 2.526 (1.062)** | 0.164 (0.811) | 0.063 (0.813) |
| High pay rate | | | | | | | -2.307 (1.273)* | -8.301 (2.996)*** |
| Unequal x high pay rate | | | | | | | | 7.666 (3.3)** |
| N | 216 | 215 | 212 | 217 | 214 | 212 | 424 | 424 |

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

### 2.1.4 Interactions and mediation on the individual level

We explored whether the effect of being assigned to a group with unequal distribution was different for participants with different gender (Table S15), because gender was a strong predictor of effort (see 2.1.3), or with different score for endorsement of the accountability principle (Table S16), because individual fairness principles might play a role in how the unequal treatment affects conservation effort. We found no evidence of such interaction effects, neither among participants with low, nor among participants with high pay rate.

We also explored whether the individual fairness judgment about one's own payment rate mediated the negative effect of the unequal condition which we found for low-paid participants

(see 2.1.3). It is plausible that the effect of disadvantaged inequality on conservation effort runs through its effect on perceived fairness. To assess this, we ran a causal mediation analysis using natural effect models following (33) with the R package *medflex* (34). For the imputation model that creates synthetic counterfactual observations, all control variables used before (see 2.1.3) were included to as best as possible control for confounding of the relationship between mediator (fairness judgment) and outcome (conservation effort). We found no evidenve for an indirect causal effect on conservation effort via an individual's fairness judgment. The natural indirect effect is small and non-significant and the mediated proportion of the total effect is close to zero (Table S17).

Note that our study was neither powered, nor designed to detect the interaction or mediation effects that were analyzed in this section. Precision of all obtained estimates is low.

**Table S15:** Results from linear mixed-effect models with village as random effect on interaction between unequal treatment and gender. The dependent variable is standardized to have mean = 0 and standard deviation = 1. With outlier adjustment.

| | Low pay rate | High pay rate |
|---|---|---|
| Variable | Model 1 | Model 2 |
| Unequal | -0.272 (0.186) | 0.024 (0.181) |
| Female | 0.677 (0.173)*** | 0.426 (0.177)** |
| Unequal x female | -0.198 (0.238) | 0.015 (0.241) |
| N | 216 | 217 |
| Marginal R^2 | 0.12 | 0.04 |
| Conditional R^2 | 0.36 | 0.35 |
| Log likelihood | -891.3 | -902.5 |

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

**Table S16:** Results from linear mixed-effect models with village as random effect on interaction between unequal treatment and endorsement of the accountability principle. The dependent variable is standardized to have mean = 0 and standard deviation = 1. With outlier adjustment.

| | Low pay rate | High pay rate |
|---|---|---|
| Variable | Model 1 | Model 2 |
| Unequal | -0.136 (0.573) | -0.257 (0.545) |
| Endorsement accountability principle | -0.01 (0.088) | -0.156 (0.088)* |
| Unequal x endorsement accountability principle | -0.141 (0.303) | 0.167 (0.28) |
| N | 214 | 217 |
| Marginal R^2 | 0.04 | 0.02 |
| Conditional R^2 | 0.24 | 0.27 |
| Log likelihood | -893.8 | -906.8 |

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

**Table S17:** Results from mediation analysis with natural effect models based on data from participants with low pay rate only. Unequal distribution is exposure variable, fairness judgment is mediator and conservation effort is outcome (unstandardized). Effect decomposition into a direct and an indirect effect (via the mediator) of exposure on outcome is shown with standard errors based on the sandwich estimator. The model includes village as a control variable. With outlier adjustment.

| Effect | Estimate | Std. Error | Pr(>\|z\|) |
|---|---|---|---|
| Natural direct effect | -6.18 | 1.509 | 0 |
| Natural indirect effect | -1.084 | 1.483 | 0.465 |
| Total effect | -7.264 | 2.049 | 0 |
| Mediated proportion (95%-CI) | 0.15 (-0.45, 0.44) | | |

### 2.1.5 Between-village variation

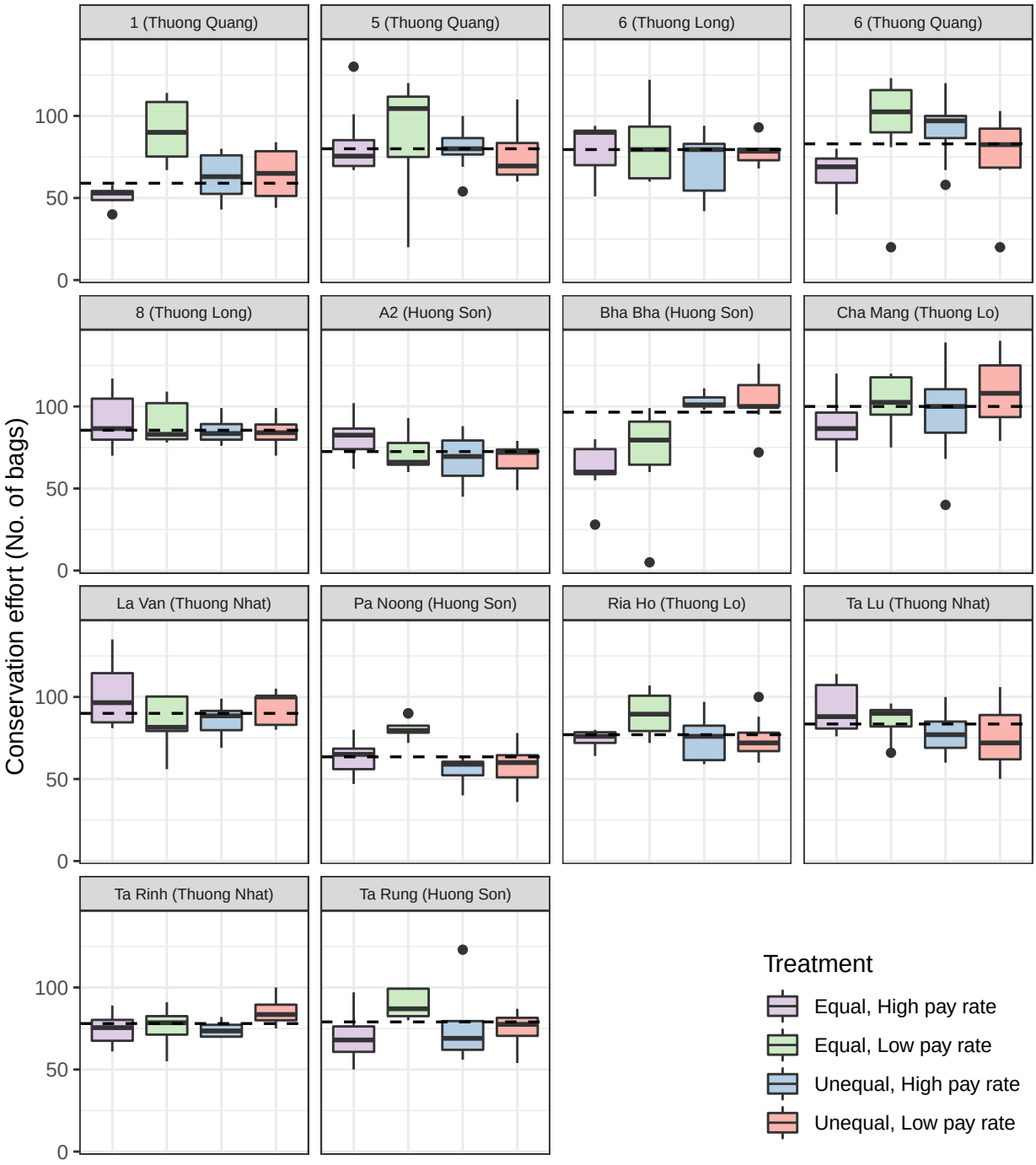


**Figure S5:** Results for all 14 experimental sites separately (village name with commune name in parentheses). Dashed horizontal lines indicate village-level median effort.

## 2.2 Conservation effort and real-world forest protection behavior

We compare conservation behavior in the real-effort experiment with real-world conservation behavior as self-reported in the survey (see 1.3). Forest protection activities like patrolling, monitoring or firefighting, which we elicited here, are explicitly part of the Vietnamese PFES program in which villages are enrolled (6), but participation on the household or individual level is de facto not enforced (65% responded being active in forest protection). When running mixed-effect models on the full dataset to predict effort in the training run and effort in the incentivized experiment with forest protection behavior as an independent variable (adjusting for treatment and the standard set of control variables as shown in 2.1.3, except the training run), there is some evidence for a correspondence between real-world and real-effort task behavior (p-values from mixed-effect models are shown in box plots, p-values from simple t-tests for the training and the experimental data are 0.007 and 0.101, respectively).

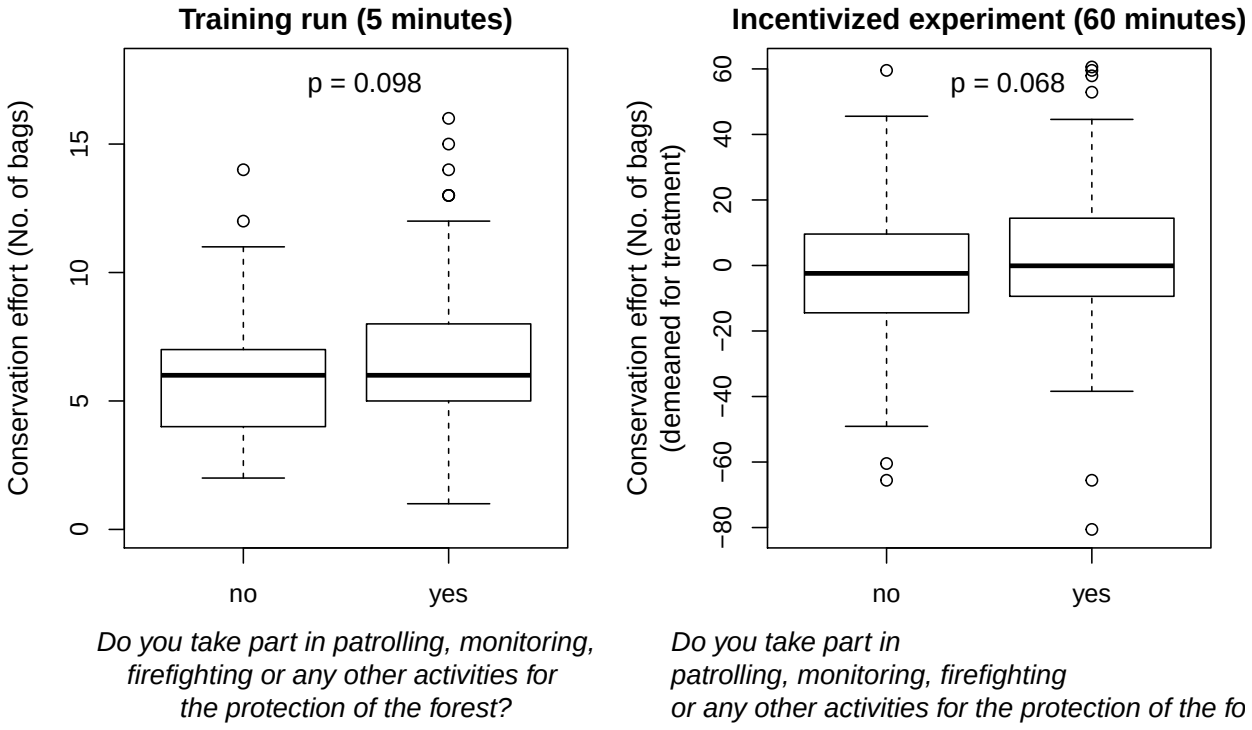


**Figure S6:** Box plots show conservation effort in the real-effort task during the training run (left) and during the incentivized experiment (right; to account for treatment effects, the plot shows observations with subtracted treatment means) for participants who report being active in forest conservation vs those who report not being active. The p-values are based on mixed-effect models.

## 2.3 Conservation effort and active working time

Time not working was measured in the survey, where participants were asked to estimate and report the number of minutes they did not spend working (see 1.3). There is a significant correlation between spending less time actively working and producing fewer bags, but also considerable unexplained variance that can be attributed to differences in working intensity (speed) rather than working minutes.

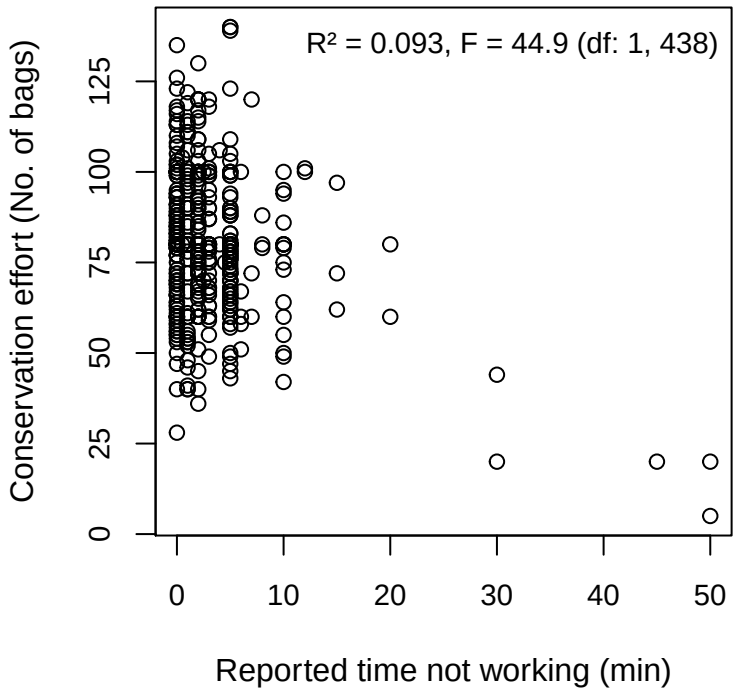


**Figure S7:** Scatter plot for time not spent working and outcome in the real-effort task. $R^2$ and F statistic from univariate linear regression are shown.

## 2.4 Effect of unequal treatment on fairness judgment

**Table S18:** Results from ordinal mixed-effect models on the effect of unequal payments on the survey measure of fairness judgment (1-5, higher values indicate higher perceived fairness of one's own pay rate) with village as a random effect. Estimated threshold coefficients and coefficients for predictors are shown with standard errors in parentheses. Coefficients for predictors are the ordered log-odds of giving a higher fairness rating.

| Variable | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| 1\|2 | -3.59 (0.299)*** | -3.137 (0.309)*** | -3.08 (0.321)*** |
| 2\|3 | -1.586 (0.181)*** | -1.05 (0.201)*** | -0.994 (0.219)*** |
| 3\|4 | -1.255 (0.174)*** | -0.695 (0.195)*** | -0.64 (0.214)*** |
| 4\|5 | 1 (0.169)*** | 1.787 (0.215)*** | 1.845 (0.234)*** |
| Unequal | -0.343 (0.18)* | -0.352 (0.182)* | -0.241 (0.252) |
| High pay rate | | 1.368 (0.194)*** | 1.484 (0.267)*** |
| Unequal x high pay rate | | | -0.231 (0.363) |
| N | 442 | 442 | 442 |
| Log likelihood | -556.428 | -529.683 | -529.48 |

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

## 2.5 Vignette results

Table S19 shows the distribution of responses for all four vignettes. The full text of the vignettes can be found in the survey questions (see 1.3). Vignettes 1, 2, 3 and 4 closely follow scenarios 1A, 1B, 1C and 1G, respectively, from (28). The results are similar to those from a US sample reported in (28), where proportions choosing the response 'Fair' were 74%, 19%, 90% and 24% in vignettes 1, 2, 3 an 4, respectively.

**Table S19:** Results for survey questions on the endorsement of the accountability principle (AP).

| Vignette | Response 'Fair' (N, %) | Response 'Unfair' (N, %) | Response in line with AP | p-value (test of proportions) |
|---|---|---|---|---|
| 1 | 237 (53.5) | 206 (46.5) | 'Fair' | 0.1408 |
| 2 | 89 (20.1) | 354 (79.9) | 'Unfair' | 0.0000 |
| 3 | 392 (88.9) | 49 (11.1) | 'Fair' | 0.0000 |
| 4 | 17 (3.8) | 426 (96.2) | 'Unfair' | 0.0000 |

## 2.6 Open-ended question on fairness of pay

After rating the fairness of their own pay rate on a five-point scale in the post-experimental survey, participants were asked *"Why and what do you think about it?"*. (see 1.3). We did

this to get a better of idea of people's perception of fairness regarding the task and treatment manipulations. Non-response was comparably high for this question (19% missing). Open-ended responses werde translated from Vietnamese to English and coded inductively. First, two authors (LL, SG) separately sampled 50 responses each and then jointly developed a coding scheme with code definitions and examples. Then, both coded all responses following this scheme, being blinded for treatment assignment and allowing for a maximum of two codes per respondent. Intercoder agreement was 84%. A discussion on differing codings led to further refinement of the code definitions. All disagreements were resolved via discussion. The final coding scheme and counts of responses are shown in Table S20. Figure S8 shows counts by treatment and fairness judgment (the five-point scale was binarized due to heavy skew towards high fairness judgments, see Figure 4B in main text). We also derived broader categories for the 12 detailed codes, dependent on what participants referred to in their response, see column "Reference for fairness judgment" in Table S20. We plot these broader categories for participants from equal and unequal groups separately in Figure S9.

**Table S20:** Coding of responses for the open-ended survey question 'Why and what do you think about it?' that followed the quantitative fairness judgment of the participant's own pay rate.

| Code no. | Response | Definition: Fairness judgment is based on... | Example | Reference for fairness judgment | N (%) |
|---|---|---|---|---|---|
| 1 | High pay (reference: experiment) | ...the comparison to other pay rate(s) of the experiment that are lower | 'Because I drew out a higher pay rate than others' | Others' pay rates in the experiment | 21 (5.3) |
| 2 | Low pay (reference: experiment) | ...the comparison to other pay rate(s) of the experiment that are higher | 'Other people get a higher pay rate' | Others' pay rates in the experiment | 23 (5.8) |
| 3 | Difference in pay rate | ...the fact that pay rates in the experiment were different, not mentioning own pay rate | 'Because of the difference in the pay rate' | Others' pay rates in the experiment | 24 (6) |
| 4 | Same pay rate | ...the fact that pay rates in the experiment were the same, not mentioning own pay rate | 'We worked together and thus the same pay rate was reasonable' | Others' pay rates in the experiment | 3 (0.8) |
| 5 | High pay (reference: unspecified or outside exp.) | ...the pay rate being perceived as generally reasonable / high without referring to experiment pay rate(s) as a comparison, but sth. else or nothing | 'It was higher than pay rate in the market' | Pay rates or total payments in general | 88 (22.1) |
| 6 | Low pay (reference: unspecified or outside exp.) | ...the pay rate being perceived as generally unreasonable / low without referring to experiment pay rate(s) as a comparison, but sth. else or nothing | 'The pay rate was low' | Pay rates or total payments in general | 41 (10.3) |
| 7 | Depends on own effort | ...the fact that total payments depended on the own effort | 'Because I received a 400 VND/piece. I will still participate with higher pay rate, e.g., 600 VND/piece' | Pay rates or total payments in general | 25 (6.3) |
| 8 | Task opportunity for money | ...the opportunity to work for income / money | 'Because I can earn some money by filling soil bags' | Pay rates or total payments in general | 12 (3) |
| 9 | Task easy/fun | ...the simplicity or joy of the activity | 'It was an easy and simple task' | Task characteristics | 88 (22.1) |
| 10 | Task useful/social/for research/for learning | ...the overall usefulness or sociality of the activity, or its value for research and learning | 'It was a useful programs to create good conditions for people to know how to make soil bags for nursery' | Task characteristics | 29 (7.3) |
| 11 | Randomness/luck | ...the fact that participants themselves were responsible for drawing their pay rate | 'Because of randomly drawing a pay rate' | Process of pay rate assignment | 14 (3.5) |
| 12 | Other | ... anything else | 'It was suitable for my health status' | - | 30 (7.5) |

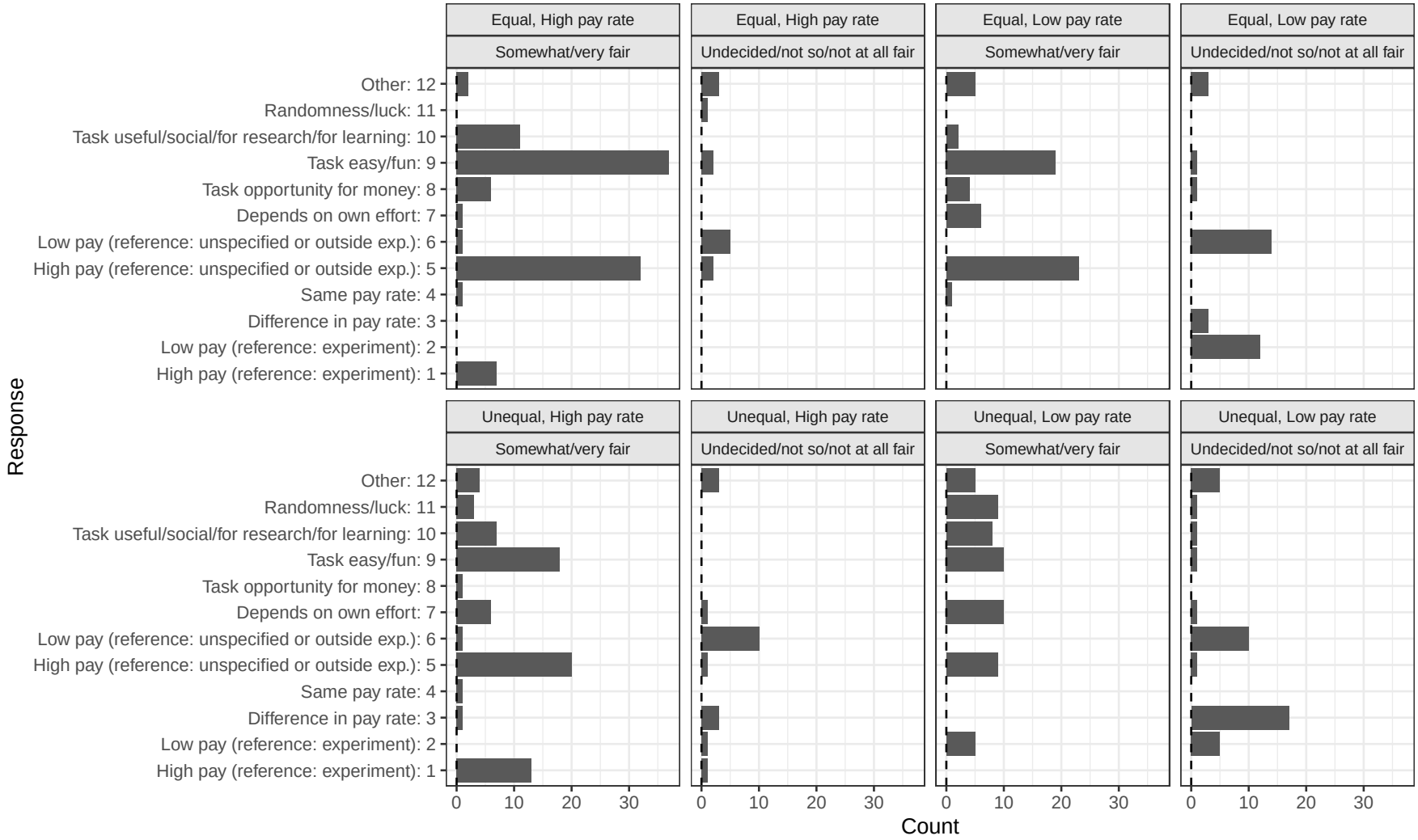


**Figure S8:** Responses for the open-ended survey question 'Why and what do you think about it?' that followed the quantitative fairness judgment of the own pay rate, split by treatment and fairness judgment (max. two codings per response were allowed).

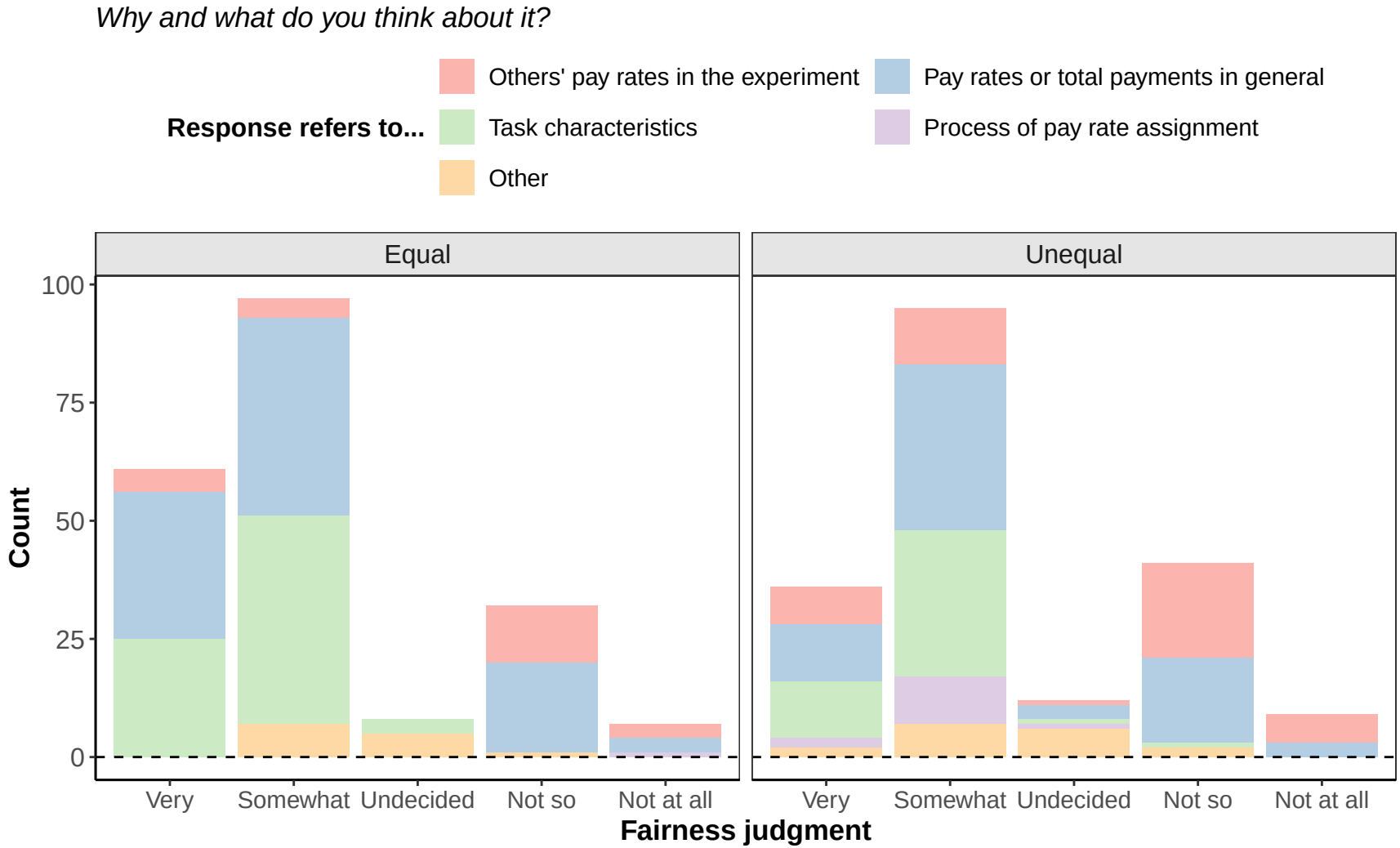


**Figure S9:** Responses for the open-ended survey question 'Why and what do you think about it?' that followed the quantitative fairness judgment of the own pay rate, categorized by what aspect of the experiment participants referred to in their response and split by equal vs unequal group and fairness judgment.

The open-ended question gave insights about how participants perceived the treatments and task (see also Figures S8, S9 and Table S20):

- Fairness judgments were often high because people thought the task was easy, sociable, useful or generally a way to generate income (code no. 8, 9, 10), e.g.: "It was an easy task, and I can make money" or: "It was interesting to work. You could learn the technique by participating and also get good cooperation with other sisters".
- In contrast, those judging payment as unfair referred almost exclusively to pay rate differences in the experiment or to differences in payments or pay rate compared to outside the experiment (code no. 1, 2, 3, 4, 5, 6, 7, 8). In both distributional treatments (equal and unequal), the reference to payment rate inequality within the experiment was highest when judging the paymen as unfair (not so fair / not fair at all).
- Low-paid participants in unequal groups gave a low fairness overwhelmingly because they did not like the pay disparities (code no. 3) or because their pay rate was too low in general (code no. 6) or compared to others (code no. 2), e.g.: "Within a working group, the pay rate should be equal" or: "It was not reasonable because some people received 400 VND/piece, some people received 600 VND/piece. I will still participate with a higher pay rate with the same productivity".
- Also among the advantaged participants in unequal groups (albeit substantially less than among disadvantaged participants in unequal groups) some thought that their pay rate was unfair due to pay disparities (code no. 3), e.g.: "Why did the pay rate differ among participants? It would be more reasonable with an equal pay rate".
- Most advantaged participants in unequal groups who thought their payment was very or somwhat fair did so because they found their pay rate was high (code no. 5), with some making explicit reference to being advantaged over others (code no. 1): "It was higher than 400 VND/piece"
- Even within the same treatment, there was considerable heterogeneity in what people thought about the absolute level of their pay rate (code no. 5, 6), also reflected in the judgment on the fairness scale. For example, both of the following responses are from high-paid participants in equal groups: "600 VND/piece was reasonable and it was higher than the market price" and "The pay rate was low. It should be 800 VND/piece. I will not participate with a lower pay rate".
- There were references to payment inequalities between participants (code no. 1, 2, 3) even in treatments with equal groups (albeit much less than in unequal groups), where

this should not be expected. Remember that equal groups did not know that another pay rate than their own existed. The most likely explanation is that some people from different groups talked to each other *after* participating in the real-effort task but *before* taking the survey. This is evident in one of the responses: "After finishing the task, I got to know that other groups got higher pay rate". We tried to prevent mixing of participants after the task, but this was difficult in some villages where the groups could not be seated very far away from each other (in any case, this could not influence the result of the incentivized experiment).

- For some disadvantaged participants, the personal accountability for their *total number of bags* (based on their own effort; code no. 7) was way more salient than the lack of accountability for their *pay rate per bag* (based on luck). They gave a high fairness rating, for example based on the following reasoning: "Because you were earning based on productivity and thus the pay rate was very good". Some people in other treatments had similar views and gave high fairness ratings with, for example, following explanations: "Doing more works will make more money" or: "It was deserved with real effort".

- While we expected and were able to show that our treatment manipulation - assigning pay rates purely based on luck rather than deservingness - violated local fairness norms of accountability, we also found deviating concpetualizations of fairness for some participants. For example, some disadvantaged participants thought that payments were very or somewhat fair because they were themselves accountable for their bad luck in drawing the envelope with their pay rate (code no. 11), giving for example the following reason: "Because of randomly drawing a pay rate". One participant even gave the following reason for a high fairness judgment (code no. 12): "Because depending on individual health". Both views are contrary to vignette results (see see 2.5), where the majority of people stated that inequalities due to factors not under undividual control (born with one hand instead of two, see 1.3), as compared to those that are under individual control (choosing to work slower, see 1.3), are unfair. Strikingly, one participant in the sample thought that gender equality in pay was unfair (code no. 6): "Because the pay rate for a man is usually different from a woman and thus this pay rate was low".

- We also found evidence in the responses that participants had opposing predictions of the effect of high absolute pay rates on conservation effort. Some stated that giving more money makes them work more: "It was not reasonable because one could receive 400 or 600 VND/piece and thus decreasing productivity. there will be more productivity with a high pay rate, 600 VND/piece because you can earn more income". Some believe that the opposite would happen, e.g.: "People who get a low pay rate try to fill the soil bags

quickly, while people with a high pay rate slow down".

## 2.7 World Values Survey data on fairness expectation

Table S21 shows World Values Survey responses to the question *"Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?"* averaged across all countries and survey waves available from the World Value Survey. Responses are on a 10-point scale from *"Would take advantage"* (= 1) to *"Try to be fair"* (= 10). The original aggregated data-file (1981-2014) was retrieved from *http://www.worldvaluessurvey.org/* in February 2020.

**Table S21:** Country-level responses to the question 'Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?' from the World Values Survey. Higher values indicate a higher expectation to be treated fairly on a scale from 1 to 10.

| Country | Average response | Sample size | Rank |
|---|---|---|---|
| Vietnam | 7.635 | 1453 | 1 |
| Uzbekistan | 7.568 | 1425 | 2 |
| Norway | 7.281 | 1021 | 3 |
| China | 7.123 | 3833 | 4 |
| Kuwait | 7.090 | 1240 | 5 |
| Rwanda | 7.082 | 3023 | 6 |
| Sweden | 7.025 | 2177 | 7 |
| Switzerland | 6.910 | 1223 | 8 |
| Taiwan | 6.838 | 2419 | 9 |
| Qatar | 6.805 | 1052 | 10 |
| Netherlands | 6.716 | 2816 | 11 |
| Indonesia | 6.666 | 1850 | 12 |
| South Korea | 6.649 | 2395 | 13 |
| Finland | 6.633 | 1010 | 14 |
| Georgia | 6.604 | 2553 | 15 |
| Philippines | 6.599 | 1200 | 16 |
| Hong Kong SAR China | 6.559 | 998 | 17 |
| Australia | 6.457 | 2858 | 18 |
| Canada | 6.424 | 2125 | 19 |
| Kyrgyzstan | 6.317 | 1498 | 20 |
| Mexico | 6.191 | 3490 | 21 |
| Mali | 6.153 | 1444 | 22 |
| France | 6.115 | 999 | 23 |
| Kazakhstan | 6.037 | 1500 | 24 |
| Lebanon | 6.027 | 1198 | 25 |
| United Kingdom | 6.015 | 1028 | 26 |
| Jordan | 6.010 | 2377 | 27 |
| Estonia | 5.995 | 1501 | 28 |
| Malaysia | 5.966 | 2501 | 29 |
| South Africa | 5.943 | 6424 | 30 |
| Pakistan | 5.923 | 1200 | 31 |
| Zambia | 5.890 | 1388 | 32 |
| Thailand | 5.867 | 2727 | 33 |
| Yemen | 5.810 | 927 | 34 |
| United States | 5.798 | 3436 | 35 |
| Palestinian Territories | 5.792 | 993 | 36 |
| Singapore | 5.735 | 1970 | 37 |
| Moldova | 5.699 | 1012 | 38 |
| Belarus | 5.667 | 1532 | 39 |
| Libya | 5.662 | 2044 | 40 |
| Iraq | 5.647 | 1187 | 41 |
| Nigeria | 5.639 | 1759 | 42 |
| Hungary | 5.611 | 996 | 43 |
| Ecuador | 5.611 | 1194 | 44 |
| Germany | 5.574 | 4053 | 45 |
| Andorra | 5.556 | 1001 | 46 |
| Russia | 5.529 | 4159 | 47 |
| Slovenia | 5.517 | 2065 | 48 |
| Italy | 5.501 | 988 | 49 |
| Ukraine | 5.494 | 2445 | 50 |
| Peru | 5.475 | 2600 | 51 |
| Uruguay | 5.465 | 1921 | 52 |
| Colombia | 5.430 | 4424 | 53 |
| Algeria | 5.415 | 1143 | 54 |
| Azerbaijan | 5.365 | 1002 | 55 |
| Tunisia | 5.352 | 1174 | 56 |
| Japan | 5.327 | 3246 | 57 |
| Egypt | 5.294 | 4534 | 58 |
| Spain | 5.269 | 2277 | 59 |
| Chile | 5.228 | 1942 | 60 |
| Turkey | 5.221 | 2905 | 61 |
| Argentina | 5.190 | 1943 | 62 |
| Armenia | 5.088 | 1051 | 63 |
| India | 4.973 | 4053 | 64 |
| Romania | 4.953 | 3105 | 65 |
| Brazil | 4.907 | 2951 | 66 |
| Ghana | 4.798 | 3061 | 67 |
| Zimbabwe | 4.744 | 1500 | 68 |
| Burkina Faso | 4.704 | 1370 | 69 |
| Iran | 4.692 | 2650 | 70 |
| Poland | 4.671 | 1829 | 71 |
| Bulgaria | 4.601 | 968 | 72 |
| Ethiopia | 4.543 | 1482 | 73 |
| Cyprus | 4.475 | 2043 | 74 |
| Morocco | 4.268 | 2284 | 75 |
| Trinidad & Tobago | 4.209 | 1990 | 76 |
| Haiti | 2.618 | 1987 | 77 |

# Supplementary Bibliography


1. G. Charness, U. Gneezy, A. Henderson, Experimental methods: Measuring effort in economics experiments. *Journal of Economic Behavior & Organization* **149**, 74–87 (2018).

2. K. M. Nelson, A. Schlüter, C. Vance, Funding conservation locally: Insights from behavioral experiments in Indonesia. *Conservation Letters* **11**, e12378 (2018).

3. U. Pascual*et al.*, Social Equity Matters in Payments for Ecosystem Services. *BioScience* **64**, 1027–1036 (2014).

4. E. H. Bulte, L. Lipper, R. Stringer, D. Zilberman, Payments for ecosystem services and poverty reduction: Concepts, issues, and empirical perspectives. *Environment and Development Economics* **13**, 245–254 (2008).

5. J. M. Alix-Garcia*et al.*, Payments for environmental services supported social capital while increasing land management. *Proceedings of the National Academy of Sciences* **115**, 7016–7021 (2018).

6. L. Loft, S. Gehrig, D. N. Le, J. Rommel, Effectiveness and equity of Payments for Ecosystem Services: Real-effort experiments with Vietnamese land users. *Land Use Policy* **86**, 218–228 (2019).

7. S. Wunder, Revisiting the concept of payments for environmental services. *Ecological Economics* **117**, 234–243 (2015).

8. J. C. Haas, L. Loft, T. T. Pham, How fair can incentive-based conservation get? The interdependence of distributional and contextual equity in Vietnam's payments for Forest Environmental Services Program. *Ecological Economics* **160**, 205–214 (2019).

9. P. McElwee, B. Huber, T. H. V. Nguyễn, Hybrid Outcomes of Payments for Ecosystem Services Policies in Vietnam: Between Theory and Practice. *Development and Change* **51**, 253–280 (2020).

10. J. Li, M. W. Feldman, S. Li, G. C. Daily, Rural household income and inequality under the Sloping Land Conversion Program in western China. *Proceedings of the National Academy of Sciences* **108**, 7721–7726 (2011).

11. J. Liu, S. Li, Z. Ouyang, C. Tam, X. Chen, Ecological and socioeconomic effects of China's policies for ecosystem services. *Proceedings of the National Academy of Sciences* **105**, 9477–9482 (2008).

12. E. Corbera, C. G. Soberanis, K. Brown, Institutional dimensions of Payments for Ecosystem Services: An analysis of Mexico's carbon forestry programme. *Ecological Economics* **68**, 743–761 (2009).

13. K. Baylis, S. Peplow, G. Rausser, L. Simon, Agri-environmental policies in the EU and United States: A comparison. *Ecological Economics* **65**, 753–764 (2008).

14. S. Uthes, B. Matzdorf, Studies on Agri-environmental Measures: A Survey of the Literature. *Environmental Management* **51**, 251–266 (2013).

15. H. Feltham, K. Park, J. Minderman, D. Goulson, Experimental evidence that wildflower strips increase pollinator visits to crops. *Ecology and Evolution* **5**, 3523–3530 (2015).

16. J. S. Kemerink-Seyoum, T. M. Tadesse, W. K. Mersha, A. E. C. Duker, C. De Fraiture, Sharing benefits or fueling conflicts? The elusive quest for organizational blue-prints in climate financed forestry projects in Ethiopia. *Global Environmental Change* **53**, 265–272 (2018).

17. K. P. Andersson*et al.*, Experimental evidence on payments for forest commons conservation. *Nature Sustainability* **1**, 128–135 (2018).

18. Ø. N. Handberg, A. Angelsen, Pay little, get little; pay more, get a little more: A framed forest experiment in Tanzania. *Ecological Economics* **156**, 454–467 (2019).

19. C. Salk, M.-C. Lopez, G. Wong, Simple Incentives and Group Dependence for Successful Payments for Ecosystem Services Programs: Evidence from an Experimental Game in Rural Lao PDR. *Conservation Letters* **10**, 414–421 (2017).

20. T. T. Gatiso, B. Vollan, R. Vimal, H. S. Kühl, If Possible, Incentivize Individuals Not Groups: Evidence from Lab-in-the-Field Experiments on Forest Conservation in Rural Uganda. *Conservation Letters* **11**, e12387 (2018).

21. F. Carlsson, O. Johansson-Stenman, P. Khanh Nam, Funding a new bridge in rural vietnam: A field experiment on social influence and default contributions. *Oxford Economic Papers* **67**, 987–1014 (2015).

22. J. Kerr, M. Vardhan, R. Jindal, Prosocial behavior and incentives: Evidence from field experiments in rural Mexico and Tanzania. *Ecological Economics* **73**, 220–227 (2012).

23. L. Loft*et al.*, Whose Equity Matters? National to Local Equity Perceptions in Vietnam's Payments for Forest Ecosystem Services Scheme. *Ecological Economics* **135**, 164–175 (2017).

24. J. He, T. Sikor, 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 (2015).

25. M. T. Bennett, China's sloping land conversion program: Institutional innovation or business as usual? *Ecological Economics* **65**, 699–711 (2008).

26. E. Uchida, J. Xu, Z. Xu, S. Rozelle, Are the poor benefiting from China's land conservation program? *Environment and Development Economics* **12**, 593–620 (2007).

27. European Comission, Agri-environment Measures: Overview on General Principles, Types of Measures, and Application. *Directorate General for Agriculture and Rural Development Unit G-4 - Evaluation of Measures applied to Agriculture* (2005).

28. J. Konow, A positive theory of economic fairness. *Journal of Economic Behavior & Organization* **31**, 13–35 (1996).

29. G. Barbato, E. M. Barini, G. Genta, R. Levi, Features and performance of some outlier detection methods. *Journal of Applied Statistics* **38**, 2133–2149 (2011).

30. R. D. Cook, Detection of Influential Observation in Linear Regression. *Technometrics* **19**, 15–18 (1977).

31. R. Nieuwenhuis, M. T. Grotenhuis, B. Pelzer, Influence.ME: Tools for Detecting Influential Data in Mixed Effects Models. *R Journal* **4**, 38–47 (2012).

32. Y. Croissant, G. Millo, others, Panel data econometrics in R: The plm package. *Journal of statistical software* **27**, 1–43 (2008).

33. T. Lange, S. Vansteelandt, M. Bekaert, A Simple Unified Approach for Estimating Natural Direct and Indirect Effects. *American Journal of Epidemiology* **176**, 190–195 (2012).

34. J. Steen, T. Loeys, B. Moerkerke, S. Vansteelandt, Medflex: An R Package for Flexible Mediation Analysis using Natural Effect Models. *Journal of Statistical Software* **76**, 1–46 (2017).