Experimental Tests of Tropical Forest Conservation Measures

Nystad Handberg, Ø. and Angelsen, A.
Norwegian University of Life Sciencie

Abstract

We conducted framed field experiments (FFEs) with local forest users in Tanzania, testing three different conservation treatments: command and control (CAC), payment for environmental services (PES) and community forest management (CFM). Our participants display more pro-social behaviour than similar studies have shown, indicating that forest specific framing is influential for participants' behaviour. We also find that treatments have strong impacts through non-pecuniary channels. CFM is as efficient as CAC in increasing pro-social forest use, despite not directly affecting the pecuniary gain. PES – as designed here – is the least effective treatment, but the results might be parameter sensitive. Women use forests more intensively than men, but are also more responsive to the treatments. The behavioural validity of the experiment is supported by strong correlation between behaviour in the experiment and stated real life forest behaviour, while treatment validity cannot be tested directly. We propose that FFEs should become a supplement to traditional impact assessments (IA) of forest conservation policies, as it avoids several challenges facing more traditional IA methods.

Key words: *framed field experiment; external validity; command and control; payment for environmental services; community forest management.*

Introduction

About one tenth of global greenhouse gas emissions originates from tropical deforestation and forest degradation (IPCC, 2013). Policy makers have therefore, under the umbrella of REDD+ (Reducing Emissions from Deforestation and forest Degradation), taken measures to reduce forest emissions. They face a problem: a myriad of possible measures exists, and the firm evidence on their effectiveness remains scant for a number of reasons. The interventions are often complex and the effects are long term (forests grow slowly). The measures may also be part of a broader political agenda by governments, aid agencies or NGOs, and a "pilot and persuade" approach might be a better strategy for project proponents than rigorous impact evaluations (Pritchett, 2008:125).

Traditional econometric methods to assess impacts of such measures face a number of challenges. Cross-sectional data suffers from unobservables affecting the performance of the forest management measures, while time-series data suffers from possible selection biases that determine the implementation of specific measures in specific areas. Randomised controlled trials (RCT) (Duflo et al., 2008) avoids selection biases by randomisation, but the method is constrained by moral, political, time and budget considerations. Intertwined effects are also difficult to observe.

Framed field experiments (FFEs) can provide a useful and efficient tool to make preliminary assessments of forest management institutions and policy measures at local level, before potentially scaling up. Economic experiments offer the appealing attribute of creating credible counterfactuals through randomisation. They allow for testing multiple treatments within the same population. Compared to laboratory experiments, used by for example Ostrom and colleagues in researching common pool resources (see Ostrom, 2006 for a review), FFEs should increase the external validity by both experimenting with the actual forest users and by framing the experiment close to (potential) real-life situations. Forests have specific attributes that call for forest specific experiments. Moral and ethical considerations, along with cultural factors, affect harvesting decisions among local forest users (FAO, 2012; Henrich et al., 2010; Levitt & List, 2007).

Key policies to contain tropical deforestation and forest degradation include command and control (CAC), payment for environmental services (PES), and community forest management (CFM) (Angelsen, 2010). These broad ideas of forest management are all founded upon solid theoretic reasoning. In order to test the effectiveness of these three policy approaches, this paper presents an experimental study undertaken with local forest users in Tanzania. We use a forest specific experimental design, building on Ostrom et al. (1994), Cardenas (2004) and Rodriguez-Sickert et al. (2008). We address three questions: What are the impacts of three different treatments (CAC, PES, CFM) on the use of a common forest stock? Do forest users, differentiated by age and gender, behave differently and do they respond differently to the treatments? What is the external validity of the experimental results, and hence the scope of the method in preliminary assessments of forest conservation measures?

The paper introduces more field relevance in the experimental design than done in the previous studies. The paper is also the first to test and compare the three forest conservation treatments. Thus, the paper aims to contribute to the method's development and the discussion on external validity, demonstrate how FFEs can be used as an *ex-ante* evaluation of forest measures, and help forest policy makers' understanding of the impact of common conservation measures on local users' behaviour.

Background and motivation for method

REDD+ and forest conservation

REDD+ can be viewed as an umbrella term for actions aimed at Reducing Emissions from Deforestation and forest Degradation, and enhance forest carbon stocks in developing countries (Angelsen, 2009). Since being launched at the international climate conference in 2007 (UNFCCC, 2007), REDD+ has generated attention and funds at global level and is seen as a relatively quick and cheap measure to limit global warming (Stern, 2006). In 2007 Norway also pledged USD 2.6 billion over five years and became the most prominent REDD+ donor, with its first bilateral agreement signed with Tanzania in April 2008.

The core idea of REDD+ is performance-based payments in a multi-level PES system: countries are paid for documented emission reductions from a future carbon market, global funds or bilateral contracts. Within countries, forest owners and users are paid for the achieved results, either individually or as a group (e.g., at village level). In practice, however, the range of policy instruments will go well beyond PES (Angelsen, 2009), and include command and control (CAC) measures, such as harvesting quotas, bans on particular use, protected areas, and land use zoning. A third major policy instrument is community forest management (CFM), understood as recognising the right and power of local communities to make decisions regarding a forest resource (Angelsen, 2009). CFM is a key element of the REDD+ strategy in Tanzania (Blomley et al., 2011; Mwakalobo et al., 2011; URT, 2012). These three core ideas in forest management are clearly different. A CAC policy aims to define and deter illegal forest use through monitoring and punishments. A PES policy seeks to change behaviour by internalising the externalities in the individual forest use decisions through economic incentives. The CFM option is a variant of the Coasian solution, where the group is given authority to manage the forest and may internalize (at least local) externalities through collective decision making.

Evaluating natural resource institutions in lab experiments

To circumvent the challenges of traditional impact assessment, laboratory experiments have increasingly been used to analyse natural resource management problems. Lab experiments were an important part of Elinor Ostrom's empirical research on managing common pool resources (CPR). Ostrom et al. (1994) claim implicitly that communication in CPR experiments is correlated with real life community management of CPRs. Their experimental finding that communication enhances aggregate welfare is used as an argument for locally founded institutions in CPR management.

Lab experiments offer several advantages (Camerer, 2011; Falk & Heckman, 2009; Ostrom, 2006): (i) the controlled environment enables researchers to identify causal *ceteris paribus* effects and to observe the effect of subtle changes in the experimental treatments (high internal validity and precision)[7]; (ii) a clear research protocol allows the experiment to be replicated; (iii) the lab is flexible to test diverse treatments with small implications for the participants, making the scope of assessment high; (iv) the experiments are often conducted at the researcher's home university, making the costs low and the practical challenges few.

The often emphasised Achilles' heel of lab experiments is their external validity (Harrison et al., 2007; Loewenstein, 1999), pointing particularly to experiments being done out of context and using participants from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) societies (Henrich et al., 2010). This remains, however, a major controversy; for instance, to what degree are behavioural responses universal and consequently valid outside the walls of the laboratory (Camerer, 2011; Levitt & List, 2007; Voors et al., 2012)

Randomised controlled trials (RCTs)

Randomised controlled trials (RCTs) offer the appealing attribute of high (local specific) external validity.[8] Since RCT treatments are implemented in real life, the actual effect on peoples' behaviour is revealed. Examples include incentivised immunization in India (Banerjee et al., 2010) and deworming pills as a measure to reduce school absence in Kenya (Miguel & Kremer, 2004). Compared with the four favourable attributes of lab experiments above (i-iv), RCTs have certain limitations. RCTs have internal validity in establishing cause-effect relationships, assuming insignificant confounding time-variant factors. However, the observed treatment effects are "crude", in the sense that only the aggregate effect is observed. In addition, replicability and scope of assessment suffer from moral, political, time, and budget constraints.

Implementing RCTs for our problem – tropical forest management – raises additional problems. Potential treatments such as CFM are to be implemented at community (village) rather than individual level, which raise the geographical scale of the experiment in ways that are challenging. Long time frames and leakages (spillover effects) are other potential problems. One of the paper authors discussed the possibility for RCT-like experiments with

[7] Internal validity refers to the ability to reveal confident cause-and-effect relationship (Loewenstein, 1999). We refer to precision as the ability to distinguish intertwined effects. A crude treatment consists of a "package" of effects where the precision is low, as only the sum of these effects is observed.

[8] Some (e.g., Deaton, 2010; Rodrik & Rosenzweig, 2010) assert that the external validity is low in RCTs, as the generalisability to other settings can be low. The term is, however, used here to describe transferability of findings in the experiment to "reality" within the participant population (Lusk et al., 2006).

REDD+ project proponents in 2008. In the end, many went for REDD+ models inspired heavily by earlier projects. Also, villages decide themselves on the use of revenues from sale of carbon credits in the voluntary carbon market, e.g., an equal payment per household or village development projects.

Framed field experiments (FFEs)

Framed field experiments (FFEs) represent a possible meeting point between lab experiments and RCTs. In the taxonomy of Harrison & List (2004), these experiments differ from lab experiments by sampling from a relevant population and by introducing a relevant commodity, task or information (framing) in the experiment, which in turn increases the external validity (Harrison & List, 2004). This is critical for non-universal findings, i.e., where culture, commodity, geography, etc. systematically affects the behaviour. FFEs also cover the middle ground between lab experiments and RCTs for the other attributes (i-iv).

FFEs on CPR management in developing countries were pioneered by Cardenas (2000, 2004). The experiments in Colombian villages have field context in its sample (local forest users) and some field context in framing (participants' payoffs are calculated from how many "months in the forest" they indicated in a payoff table: from one to eight months). Also, the experiments are held where the participants make their real life decisions. The results indicate that communication increases cooperation in the group more than external regulations, making the case for CFM as a forest conservation tool.

In similar experiments from rural Colombia, Rodriguez-Sickert et al. (2008) find that cooperation among participants increase when a norm of only extracting one of eight tokens each round is introduced, together with a moral speech of the importance of the norm. Cooperation increase further when a positive probability (20%) of punishment of norm-breakers is introduced; cooperation is also responsive to increased levels of punishment.

Table 1 compares "harvest rates" (a fraction of the maximum possible token extraction each round) at the experiment level in the three seminal CPR experiment studies discussed. Comparing the harvest rates should be done with care, since there are numerous differences between the experimental designs (parameters, sample, instructions, level of scrutiny, framing, etc.). Still, the studies share many similarities and some of the large differences in harvest rates are noteworthy

Table **24Table 24 Rough summary of constructed harvest rates in three similar CPR experiments**

Study	OA	Treatments	Difference	N (experiments)
Ostrom et al. (1994:154)a	0.76	0.46	0.30	4
Cardenas (2000:316)b	0.55	0.47	0.08	10
Rodriguez-Sickert et al. (2008:220)c	0.58	0.29/0.34/0.46	0.29/0.24/0.11	14/26/16

[a] The participants invested in individual wealth at a fixed rate or in a CPR that had a concave return function (the more participants invested in the CPR, the less return). The reported "harvest rates" are calculated from the attained yield as a percentage of maximum possible yield. The treatment is communication.

[b] The reported "harvest rates" are calculated from absolute harvest decisions. The treatment is communication.

[c] The reported "harvest rates" are calculated from absolute harvest decisions. The treatment is command and control with high fine/low fine/rejected fine. In the rejected fine category, the participants were offered, but refused the sanctioning mechanism.

The harvesting rates under OA in the are similar, but lower than the harvest rate bottom two (Colombia) studies of Table 1 under OA in the Ostrom et al. study. The main difference between them is framing and sample. Whereas Ostrom et al. (1994) conducted abstract CPR experiments with students, Cardenas (2000) and Rodriguez-Sickert et al. (2008) introduced some field context and conducted the FFEs with rural Colombians. If the difference is due to the framing, introducing even more field context should further increase the pro-social behaviour in the OA situation. The norm treatment of Rodriguez-Sickert et al. (2008) implies an intended harvest rate of 0.125 (1/8). The strong punishment makes the expected payoff of complying with the norm clearly higher than violating it. Also, the participants are exposed to a short moral speech of the benefits of complying with the norm. Still, the harvest rate under the treatment is 0.29, revealing that there are substantial violation of the norm.

Conceptual framework

Our conceptual framework builds on Levitt & List (2007). In the open access (OA) situation, represented by Eq. [1], a utility-maximising individual *i* makes a harvesting decision *a*, based on two considerations: *a*'s effects on individual wealth *W* and *a*'s effect on non-pecuniary moral payoffs *M*. *W* depends on the decision *a* and the stakes involved *v*. *M* depends on norms affecting forest decisions *n* and the level of scrutiny *s*, in addition to *a* and *v*. In addition, Ferraro & Price (2013) suggest that a vector of individual specific characteristics $\theta$is important in determining both *W* and *M*. For our use, the important attribute of the model is distinguishing the two channels, *M* and *W*, and not creating an exhaustive list of factors determining the two channels.

OA $U_i(a,v,n,s;\theta) = W_i(a,v;\theta) + M_i(a,v,n,s;\theta)$

Introducing command and control (CAC) *p* has a direct negative effect on *W* in Eq. [2], as decision *a* is constrained by a law and violators are exposed to punishment. CAC could also affect *M*. Rodriguez-Sickert et al. (2008:226) suggest that low fines might be analogous to yellow cards on the football field: "low fines and yellow cards may sometimes stabilize norm compliance in a world of feeble social order." This potential effect is represented by δ>1.

CAC $U_i(a,v,n,s,p;\theta) = W_i(a,v,p;\theta) + \delta_i M_i(a,v,n,s;\theta)$

In contrast, introducing a payment for environmental services (PES) *b* has a positive effect on *W* in Eq. [3], as individual *i* also will be rewarded for not harvesting trees. However, PES might also negatively affect *M*. Gneezy & Rustichini (2000) and others (Bowles, 2008; Muradian et al., 2013) suggest that introducing monetary payments can crowd out intrinsic motivations, which in our case would decrease the magnitude of *M*, as represented by 0< γ<1.

PES $U_i(a,v,n,s,b;\theta) = W_i(a,v,b;\theta) + \gamma_i M_i(a,v,n,s;\theta)$

Lastly, introducing community forest management (CFM) decentralises power to the local level, which in turn increases incentives for scrutinising others and complying with norms (Ostrom et al., 1994). In our framework, this implies an increased importance of *M*, as represented by α>1. Notably, in our framework CFM does not directly affect *W*.

CFM $U_i(a,v,n,s;\theta) = W_i(a,v;\theta) + \alpha_i M_i(a,v,n,s;\theta)$

Field work, experiments and methods

Study area and sampling

Tanzania was a suitable choice for our field work: it relies heavily on its forest resources (TNRF, 2009; World Bank, 2008), it is one of the most active REDD+ countries on the continent (Angelsen, 2009; Blomley & Iddi, 2009), and it experiences one of the highest absolute losses of forest area in Africa (FAO, 2011). We conducted the experiments in three regions – Shinyanga, Singida and

Morogoro – located in the North-West,

centre and South-East of the country to capture some geographical variations, which we believe makes our results more general. Seven villages within the three regions were selected in collaboration with district authorities or local level NGOs. An absolute criterion in the village selection was an accessible forest within walking distance of the village centre (roughly 5 km). Beyond forest proximity, villages have considerable variation in distance to the nearest city, accessibility, livelihoods (as indicated by the variation in number of livestock), population, forest ownership and externally driven forest conservation projects, as shown in We constructed a list of all village households from the official village records, and 40 households were randomly drawn from the list. From each selected household, the man or the woman was randomly selected (by a coin-flip) and placed in groups of eight participants. In total, 36 experiments were conducted, giving an overall sample of 288 participants. Randomisation at both stages (selecting participants and selecting treatment groups) avoids possible randomisation bias (Harrison et al., 2009; Heckman & Smith, 1995).

The experiment consisted of two parts with six rounds in each part. In every round, each participant privately decided how many tokens (trees) to withdraw from the stock. The participants were constrained by a technical harvest limit of 5 trees per round. After all eight participants had made their decisions, the aggregated number of removed tokens was announced to the group. Before the start of the next round, the stock grew by two tokens for every ten remaining tokens, up to a maximum stock size of 160 tokens. The participants' payoffs are consequently given by Equations [5, 6, 7].

Table 25 Variation in village characteristics

Table 25 Variation in village characteristics

Village	District	Region	Distance to forest frontier[a]	Distance to nearest town[a]	Public transport	Total number of livestock[b]	Population	Ownership of forest	Externally driven conservation projects
Mughunga	Singida Rural	Singida	3	45	1/day	1050	1922	Reserve and community managed	Community conservation
Nduamughanga	Singida Rural	Singida	2	67	1/day	1834	2970	Reserve and community managed	Community conservation
Busongo	Kishapu	Shinyanga	3	50	No	1850	2219	Ngitilis[c]	None
Ngulu	Kahama	Shinyanga	2	77	No	1265	2234	Ngitilis[c]	REDD piloting
Zombo	Kilosa	Morogoro	5	18	Several times/day	15	3401	Community managed	None
Dodoma-Isanga	Kilosa	Morogoro	1	35	No	52	1308	Community managed	REDD piloting
Muhungamkola	Morogoro Rural	Morogoro	4	45	No	1735	1756	Reserve and community managed	None

[a]numbers in kilometres; [b]cows and donkeys; [c]Swahili word for "enclosure", a traditional arrangement where private or communal areas are managed as woodland or pasture (Blomley & Iddi, 2009). Sources: The village's chairperson, the district's forest officer, or the area's forest conservation project coordinator.

While we do not pursue an analysis of, for example, differences in responses based on these characteristics, the variation enhances the overall robustness of the findings.

Table 25 Variation	Externally driven conservation projects	Community conservation	Community conservation	None	REDD piloting	None	REDD piloting	None	[a] numbers in kilometres; [b] cows

We constructed a list of all village households from the official village records, and 40 households were randomly drawn from the list. From each selected household, the man or the woman was randomly selected (by a coin-flip) and placed in groups of eight participants[9]. In total, 36 experiments were conducted, giving an overall sample of 288 participants.[10] Randomisation at both stages (selecting participants and selecting treatment groups) avoids possible randomisation bias (Harrison et al., 2009; Heckman & Smith, 1995).[11]

The experiment consisted of two parts with six rounds in each part.[12] In every round, each participant privately decided how many tokens (trees) to withdraw from the stock. The participants were constrained by a technical harvest limit of 5 trees per round.[13] After all eight participants had made their decisions, the aggregated number of removed tokens was announced to the group. Before the start of the next round, the stock grew by two tokens for every ten remaining tokens, up to a maximum stock size of 160 tokens. The participants' payoffs are consequently given by Equations [5, 6, 7].

[9] From precedent set by Ostrom et al. (1994:108), with the justification that it "approximates some of the characteristics of larger groups or conflict-ridden small groups".

[10] All participants completed the experiment, but one participant (clearly intoxicated) went missing after the experiment and before the interview, meaning that most individual data for this participant is missing.

[11] Participation in our experiment was so popular that we never encountered problems in randomly selected participants not appearing. Therefore, the effect of self-selection should be small.

[12] The number of rounds was not announced to the participants to prevent backward induction. If the question arose, the participants were told that the number of rounds depend on their behaviour (which is true, as the experiment ends with stock depletion).

[13] The technical harvest limit decreases with stock size, as shown in Eq. [6]. This information is provided to the participants through the "maximum harvest table" presented in the appendix.

Village	District	Region	Distance to forest frontier[a]	Distance to nearest town[a]	Public transport	Total number of livestock[a]	Population	Ownership of forest
Mughunga	Singida Rural	Singida	3	45	1/day	1050	1922	Reserve and community managed
Nduamughanga	Singida Rural	Singida	2	67	1/day	1834	2970	Reserve and community managed
Busongo	Kishapu	Shinyanga	3	50	No	1850	2219	Ngitilis[c]
Ngulu	Kahama	Shinyanga	2	77	No	1265	2234	Ngitilis[c]
Zombo	Kilosa	Morogoro	5	18	Several times/day	15	3401	Community managed
Dodoma-Isanga	Kilosa	Morogoro	1	35	No	52	1308	Community managed
Muhungamkola	Morogoro Rural	Morogoro	4	45	No	1735	1756	Reserve and community managed

[5]

$$x_{it} < \bar{x}_{it} = \begin{cases} 0 & if \quad S_t < 8 \\ 1 & if \quad 8 \leq S_t < 16 \\ 2 & if \quad 16 \leq S_t < 24 \\ 3 & if \quad 24 \leq S_t < 32 \\ 4 & if \quad 32 \leq S_t < 40 \\ 5 & if \quad 40 \leq S_t \leq 160 \end{cases}$$

$$S_t = \left(S_{t-1} - \sum x_{jt}\right) + \beta$$

$$\beta = \begin{cases} 0 & if \quad \left(S_{t-1} - \sum x_{jt-1}\right) < 10 \\ 2 & if \quad 10 \leq \left(S_{t-1} - \sum x_{it-1}\right) < 20 \\ 4 & if \quad 20 \leq \left(S_{t-1} - \sum x_{it-1}\right) < 30 \\ 6 & if \quad 30 \leq \left(S_{t-1} - \sum x_{it-1}\right) < 40 \\ \vdots & \vdots \qquad \vdots \\ 26 & if \quad 130 \leq \left(S_{t-1} - \sum x_{it-1}\right) \leq 134 \\ 160 - \left(S_{t-1} - \sum x_{it-1}\right) & if \quad 134 < \left(S_{t-1} - \sum x_{it-1}\right) \leq 160 \end{cases}$$

Eq. [5] gives participant *i*'s payoff in round *t*, *yit* as the product of the participant's harvest in the particular round, *xit* and the price per token (TZS 100). Eq. [6] restricts the harvest (technical harvest limit) by the size of the stock in the given round, *St*, both excluding the possibility of a negative stock and adding a realistic element with tree harvesting becoming more difficult with higher scarcity. Eq. [7a, 7b] give the stock size as a function of the aggregated harvest in the previous rounds and the forest growth, *β*. Together, these mechanisms simulate the attributes of a CPR and construct the social dilemma for the participants: maximise own payoff or maximise the group's payoff.

Due to Eq. [6], a relative harvest rate (0-1) is constructed as the ratio between actual harvest and maximum possible harvest. For example, five trees harvested from a stock of 80 trees and three trees harvested from a stock of 25 trees constitute the same harvest rate: 1. The experiment characterises deforestation and forest degradation as harvest rates above the sustainable level, i.e., the level that maintains the forest stock. The strictly sustainable harvest rate for a group in the first rounds of the two parts of the experiment is 0.3. Sustainable harvest rates in other rounds depend on previous rounds' harvesting.

Framing

In addition to a relevant sample, the study focuses on field context in three forms: in setting, in commodity and in task. The aim of adding field relevance to the experimental design is to increase the external validity of the experiment (Harrison & List, 2004); factors such as the priming effect (Bargh, 2006) could make the participant include relevant (non-monetary) considerations in their experiment decisions.

The experiment is set in the participants' villages, as in Cardenas (2004) and Rodriguez-Sickert et al. (2008). Allowing the participants to make their decisions close to where they normally take their forest decisions, increases the field relevance of the setting. Moreover, familiar surroundings for the participants reduce possible noise arising from alien labs (Harrison & List, 2004). The previous studies used complicated payoff tables of 11x28 cells (Cardenas, 2004; Rodriguez-Sickert et al., 2008) and 5x11 cells (Ostrom et al., 1994) as the commodity through which the participants reveal their decisions. Although the commodities in this experiment are clearly not real trees, the participants recognise them as pictures of trees – which could affect their choices (and increase external validity). The experiment thus has some field context in commodity. The possible priming effect found in the picturing of other concepts, such as money (Vohs et al., 2006) and old age (Bargh et al., 1996), could also exists in replicates of trees. Lastly, instead of making the participants fill in their decisions in the payoff tables (as in the previously mentioned studies); our participants had to physically remove the cardboard trees. The task has field context, as the act demands some physically effort. Filling in forms reduces the trees to numbers on a sheet of paper, and may also exclude illiterate forest users.

The treatments

The first part (rounds 1-6) was identical for all participants and played as open access (OA), as described above. For the second part (rounds 7-12), the stock size was reset to 80 tokens, and one of four treatments was introduced: command and control (CAC) payment for environmental services (PES); communication, simulating community forest management (CFM); and a continuation of the open access (OA) situation. While the sequence of treatments in a given village is random, the distribution of treatments between villages is uneven (small sample size) (Table 26). We therefore control for village fixed effects in the empirical analysis.

Table 26 Distribution of experiments by villages and treatments

Treatment \ Village	1	2	3	4	5	6	7	Total
OA	2	0	0	0	1	1	2	6
CAC	1	1	2	2	2	1	1	10
PES	1	3	1	2	1	2	0	10
CFM	1	2	2	1	1	1	2	10

In the CAC treatment, a rule that the participants can legally harvest 0, 1, 2 or 3 trees was announced. The participants were told that it is possible to harvest more, but that the technical limit still applies. After each decision, the researcher threw a die. If the die showed a five or a six, the researcher inspected the harvest decision of the participant. If the rule was broken, the participant was penalised with ten trees, in addition to losing all of the trees harvested in that round (but the trees were removed from the stock).

In the PES treatment, the participants were paid TZS 80 for the trees they decided *not* to harvest, as well as the TZS 100 for the ones they harvested. The participants as such received $100x+80(z-x)$ in payoffs, where x is the number of harvested trees and z is the technical limit. The payment for not harvesting was set at 80% of the reward for harvesting, as setting the payment equal to or higher than the harvest-reward should be trivial (in theory); the participants would choose to receive the payment. This may also be more realistic for a possible PES system, i.e., full compensation is not given (see later discussion).

In the CFM treatment, the participants were allowed to communicate collectively and in privacy for three minutes at the start of each round (cheap talk). The instructions stated that they could talk about harvesting decisions, but gave no further guidelines. After the three minutes session, the participants were again prohibited from communicating until the round was completed. The actual individual harvesting decisions were still unobservable to the other members of the group

In the OA treatment, the participants were exposed to an identical experimental design as in the first part. In this way, possible learning effects can be controlled for. The experiment thus has a mixed-design, where tests are made both within treatment groups and relative to the OA treatment group.

Results and discussion

Impact of treatments

The evolving harvest rates through the twelve rounds are shown in Figure 5. In the first part (OA situation), the mean harvest rates decrease, but stabilise at 0.35-0.45 in the last two rounds for the four treatments. In the second part, the means differ more. The mean in the OA situation of part two remains in the same interval as in the first part, while the means for the other treatments decrease. The PES treatment follows a similar pattern as the OA situation, although with lower harvest rates. The CFM treatment induces the lowest harvest rates; while the harvest rates in CAC treatment tend to be between the PES and CFM treatments.

Figure 5 Mean harvest rates through the 12 rounds, impact of treatments in part 2. Table 27 aggregates and compares the mean harvest rates in the two parts of the experiments by the four treatments. The decrease in aggregated mean harvest rate in the open access situation from part one to part two is insignificant. This supports the internal validity of the experiment, as the two parts in this treatment are identical by design. Furthermore, the CFM treatment decreases the mean harvest rate by 25 pp (percentage points), the CAC treatment decreases the mean harvest by 18 pp, and the PES treatment decreases the mean harvest rate by 10 pp. The impacts of the three treatments are significant.

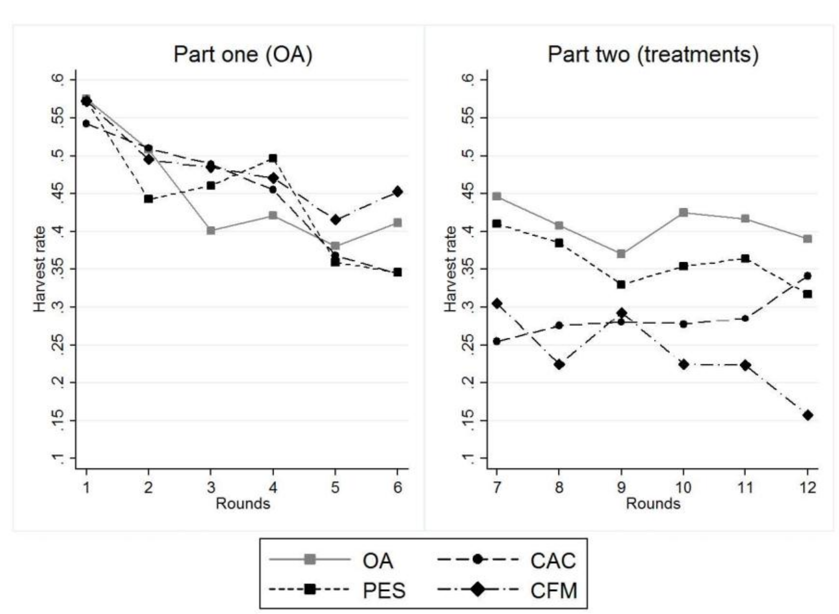

Table 27 T-test comparing the mean harvest rates in each experiment's part one and two by treatment

Treatment	Harvest rate in 1st part	Harvest rate in 2nd part	Difference (2nd part–1st part)	N
Open access	0.453	0.41	-0.043	6
	(0.03)	(0.03)	(0.04)	
CAC	0.462	0.29	-0.177***	10
	(0.03)	(0.02)	(0.04)	
PES	0.457	0.362	-0.095**	10
	(0.03)	(0.02)	(0.04)	
CFM	0.486	0.238	-0.248***	10
	(0.03)	(0.02)	(0.04)	

Difference between the means is stated in column four Standard errors in parentheses. ***,**,*: significant at the 1, 5 or 10% level

The harvest rates in the first part of the experiment are clearly lower than the "OA harvest rate" of Ostrom et al. (1994), and lower than the harvest rates of Cardenas (2000) and Rodriguez-Sickert (2008). This indicates that more framing (i.e., realism) in the experimental design increases non-pecuniary considerations among the participants and hence increases pro-social behaviour.

As confounding factors at the experiment level or village level can affect the results the treatment effect should be robustness-tested.

Table 28 regresses experiment harvest rates on the three treatments (as a categorical variable, relative to the OA situation) and three experiment level control variables: *Sequence* (the order of experiment conduction in each village), *Femgroup* (the number of female participants in each experiment) and *Agegroup* (the mean age of each experiment). Model (1) has the aggregated mean harvest rate in part two of each experiment as the dependent variable, while Model (2) has the difference in aggregated mean harvest rate between parts one and two (part2-part1) as the dependent variable. Both models include village fixed effects (not reported).

Table 28 Regressing aggregated harvest rates on treatments and control variables

	(1)Harvest rates in part two	(2) Change in harvest rates
CAC	-22.163***	-28.929***
	(6.54)	(8.82)
PES	-8.734	-13.568
	(6.46)	(8.72)
CFM	-21.316***	-29.804***
	(6.28)	(8.47)
Sequence	4.018**	0.376
	(1.53)	(2.06)
Femgroup	1.188	2.342
	(1.33)	(1.79)
Agegroup	0.910*	1.190*
	(0.47)	(0.64)
Constant	-7.876	-56.367*
	(22.80)	(30.76)
R^2	0.517	0.445
N	36	36

Dependent variables stated in row one. Both models include village fixed effects (not reported).Standard errors in parentheses .***,**,*: significant at the 1, 5 or 10% level

When we control for village fixed effects, include control variables at the experiment level and measure the treatment effects relative to the OA situation, the CAC treatment has the largest impact on aggregated mean harvest rates (22 pp), followed by the CFM treatment (21 pp) with both effects being significant. The PES treatment has an insignificant impact on aggregated harvest rates. When controlling for differences in harvest rates in the first part of the experiment (model 2), the CFM treatment has a larger impact than the CAC treatment, but the differences are insignificant.

All groups are exposed to open access in the first part. The second column of Table 27 indicates that the internal validity is high, as the mean harvest rates in the four treatment groups are insignificantly different. The left side of Figure 5 supports this claim and furthermore indicate that some learning effects are present. The harvest rates decline and stabilise in the last two rounds. This stabilised trend continues through the OA treatment in the second part of the experiment. Thus, the experiment has the necessary number of rounds in the first part for the participants to settle at stable harvesting. Further, since the number of rounds is unknown to the participants, we did not observe any "end-of game" effects, in the form of maximum harvest rates in the last round.

The harvest rates under open access are higher than the sustainable harvest rate, but are notably lower than the three studies of (0.76, 0.55, and 0.58). This indicates that M from Eq. [1] has a clear impact on harvest decisions, more so than the previous studies with less framing and one with a non-relevant sample. Even without a specific forest conservation treatment the villagers in our study area demonstrate fairly strong pro-social behaviour. In the CAC treatment, the rule deterring the participants from harvesting more than three trees implies a legal (and endorsed) harvest rate of 0.6 in the first round of the second part. The observed harvest rate is still below 0.6, implying that the treatment is more effective than intended. Related to our conceptual framework, δhas a positive effect and enhances moral considerations, e.g., in the form of norm compliance (Rodriguez-Sickert et al., 2008). That our CAC treatment has a greater impact on moral considerations than the treatment of Rodriguez-Sickert et al. could be due to the specific framing of our experiment.

Of the three treatments, PES has the least impact on harvest rates. The small impact might be explained by the introduction of only partial (80%) compensation for non-harvesting, and /or crowding out of intrinsic motivation that reduces the effectiveness of financial incentives (indicating a substantial effect of γin Eq. [3]). While the partial compensation may be the most realistic representation of a possible PES scheme, it could violate Gneezy's & Rustichini's (2000) recommending title: "Pay enough or don't pay at all". Further experiments with different levels of payments would provide a more complete test of their hypothesis in our context.

The CFM treatment has the largest impact on harvest rates (Table 27). The finding of a large pro-social effect of communication (strong effect of αin Eq. [4]) is consistent with the literature, cf. the comprehensive review by Ostrom (2006). One caveat, however, can be highlighted by an anecdote from this field study. In one village, an individual with a leading position in the village was randomly selected as a participant in a CFM experiment. As might be expected, the leader took charge when communication was allowed. When observing the harvest decisions, all participants – except one – had low and similar harvest rates. The exception was the village leader, who harvested at near maximum. He consequently received one of the highest payments throughout the study.

Such elite capture is according to Brockington (2007) a severe stumbling block for the success of CFM measures in Tanzania. The ability of FFEs to capture this issue is discussed further in subsection 0.

Individual specific characteristics

To the extent individual characteristics are predictive for forest use, understanding these relationships can be helpful for conservation planning. We regress individual mean harvest rates throughout the experiment on gender and age, as well as on stated forest use (experiment fixed effects. The models

Table 29 Individual characteristics as predictors for individual mean harvest rates

	(1) Village fixed effects	(2) Experiment fixed effects
Gender (female=1)	7.116**	7.435***
	(2.99)	(2.86)
Age	-1.441***	-1.325**
	(0.54)	(0.53)
Age^2	0.017***	0.016***
	(0.01)	(0.01)
Relative forest use (middle third relative to bottom third)	10.331*** (2.92)	10.195*** (2.90)
Relative forest use (upper third relative to bottom third)	22.555*** (6.10)	23.407*** (5.98)
Absolute forest use	1.707***	1.071*
	(0.59)	(0.58)
Commercial forest use	10.451***	10.392***
	(3.74)	(3.68)
Constant	12.104	49.322***
	(21.79)	(12.37)
R^2	0.277	0.218
N	287	287

Dependent variables: individual mean harvest rate (pct.) in both parts of the experiment.
Controls in both models (not reported): variables on stated household land and stated paid work (wealth indicators). Controls in model (1): *treatments*, *Sequence*, *Femgroup*, and *Agegroup*.
Standard errors in parentheses.
***,**,*: significant at 1, 5 or 10% level.

Women have a significantly higher mean harvest rate than men (c. 7 pp). Also, age and mean harvest rates are negatively correlated (c. -1.4 pp per year), but the relationship is U-shaped with 42 years as the minimum point. The on stated forest use are discussed in the next subsection.

The gender effect is in contrast to the eco-feminism literature (in particular Shiva (1989)), which suggests that women are "closer to nature than men" and accordingly are better conservationists (Agarwal 2010:41). In the experimental literature, however, the results are mixed (Brown-Kruse & Hummels, 1993; Cardenas et al., 2011; Schwieren & Sutter, 2008).[14]

Several experimental studies have found a positive relationship between age and pro-social behaviour (Carpenter et al., 2008; Grossmann et al., 2010; Meier & Frey, 2004). Grossman et al. (2010) even suggest that this robust finding should make older individuals hold key decision making roles. Since the finding of this study suggests that individuals are becoming more pro-social until the age of 42 and less pro-social from then on, the view of Grossman et al. is not supported by this study.

In exploring the impacts of the treatments by gender (Table 30), we find that the

[14] The gender result could also be driven by lower income levels among women (making the incentives for high harvest rates higher) and/or that women use forest resources more intensively in their daily life. The variables on stated forest use, stated land holdings and whether family members have paid off-farm work attempts to control for these potential confounding factors.

treatments mainly have an effect through women. CAC is the only treatment that relative to the OA setting produces a significant decrease in harvesting among male participants (-13 pp), and the effect is smaller than the overall effect on participant behaviour. Both the CAC and CFM treatments have strong and significant impacts on harvesting among female participants (-25 pp and -28 pp).

Table 30 Interaction effect of treatment and gender

	Individual harvest rate
Gender (female=1)	23.206*** (6.88)
Age	-1.296** (0.53)
Age2	0.015*** (0.01)
Relative forest use	11.226*** (2.40)
Absolute forest use	1.470*** (0.56)
Commercial forest use	10.519*** (3.73)
Agegroup	1.347*** (0.42)
Femgroup	0.982 (1.26)
CAC	-12.877** (6.31)
PES	-1.723 (6.06)
CFM	-7.207 (6.14)
CAC*gender	-24.528*** (8.57)
PES*gender	-12.931 (8.64)
CFM*gender	-28.261*** (8.67)
Constant	15.225 (21.75)
R2	0.345
N	287

Dependent variable: individual mean harvest rate (pct.) in part two of the experiment. Control variable: *Sequence*, village fixed effects, variables on stated household land and stated paid work (wealth

Table 29). A commercial forest user has a c. 10.4 pp higher harvest rate than a noncommercial forest user, another weekly trip in the forest to collect forest products (absolute forest use) implies a 1-1.7 pp higher harvest rate, and participants who state to be in the top (middle) third of relative forest product use in the village harvest 23 pp (10 pp) more than the bottom third. The enumerator who observed the participants' harvesting decisions noted that participants often picked cardboard

External validity: behavioural validity and treatment validity

The findings are more interesting and policy relevant if the experimental design possesses external validity. We introduce a decomposition of external validity into "behavioural validity" and "treatment validity". The former concerns the correlation between the participants' behaviour in the experiment and their real life behaviour, while the latter concerns the correlation between the experimental treatments and the real life policy measures they aim to simulate.

The behavioural validity of the experiment is supported by the strong and significant correlation between harvest rates in the experiment and stated real life forest harvesting (

The treatment validity of the experiment is not observable in the current data set and is consequently left for discussion. The aim of the specified treatment is not to emulate specific and complex forest conservation measures, but rather to explore the effect of the intended mechanisms of three core ideas within forest conservation: deterrence, economic incentives and decentralisation. Still, if the impacts of the treatments work through irrelevant factors or if the impacts are parameter sensitive, treatment validity is threatened.

The parameters in the CAC treatment (1/3 chance of being audited and a penalty of 10 trees) are strong, but not unrealistic. The results show the effect of the treatment is stronger than intended (harvest rate < 0.6). Therefore, lowering the parameters (assuming linearity) should still be expected to produce strong treatment impacts.

Potential irrelevant factors affecting the CAC treatment impact include Hawthorne effects and "white-man treatment effects" (Cilliers et al., 2012). The Hawthorne effect, that participants' behaviour is altered merely by the scrutiny of researchers, and the bolstering of such an effect by the presence of a white researcher,[15] could be stronger in this treatment than the others, as the (white) researcher is more involved. This could explain the strong impact of the treatment. Notably though, real life implementation of a CAC measure would by definition also involve scrutiny of forest users (potentially also by foreign donors and NGOs).

The parameters in the PES treatment (80% individual compensation for the trees not harvested) are specific in that collective payment, payments conditional on remaining forest stock or full (and other levels of) compensation are not considered. The difference in payment level and distribution mechanisms might produce different impacts, and the results should be regarded specific to this payment scheme. Generalisation to other PES schemes is therefore premature, and calls for more experimental research. The parameters in the CFM treatment (3 minutes of collective communication at the start of each round) are realistic in that the participants are allowed to agree upon collective strategies and follow up on the aggregated results throughout the experiment. Communication as a CPR experiment treatment is the most widely used treatment of this experiment's three treatments (Ostrom and Cardenas in particular). Variants with longer communication time and less or no repetition have also found to have strong pro-social effect on participants' behaviour. Therefore, the treatment's impact should not be particularly parameter sensitive.

The impact of the CFM treatment might, however, be contaminated by other factors than those merely related to the core idea of CFM. In particular, the scope of elite capture may be different in the experiment treatment than in a real life CFM situation. There is no heterogeneity in the endowed forest stock participants' choice set, the small group size increases scrutiny by others, and the payoffs of cheating other are relatively low. The lack of scope for this type of anti-social behaviour could create a pro-social bias of the treatment.

In sum, the external validity of the experiment is supported by its high behavioural validity, while its treatment validity is left for discussion. Despite all three treatments possible being parameter sensitive and other causal factors potentially affecting behaviour than intended, there are strong arguments for the treatments capturing and distinguishing three core ideas within forest conservation management. To further assess the treatment validity, tests could be conducted. Identical experiments with varying treatment parameters can be conducted to test the robustness of the findings, as can sensitivity to group sizes.

Conclusions

Our results suggest that the behaviour of local Tanzanian forest users is not merely a result of individual wealth considerations, but also moral, non-pecuniary considerations. The harvest rates are well below selfish payoff-maximising behaviour under open access, and lower than harvesting rates in previous and less framed experimental studies. The more pro-social behaviour points to the relevance of experiment sample and framing in FFE as compared to lab experiments. Policy makers within the forest conservation and REDD+

[15] Cilliers et al. (2012) found that participants of dictator games in Sierra Leone significantly increased their pro-social behaviour by substituting a black Sierra Leonean researcher with a white American researcher.

community should not ignore the non-pecuniary considerations local forest users make.

Among the three treatments, CAC and CFM have strong and significant impacts on forest use. The strong effect of CFM (communication) is in line with other findings and underscores how local resource control can lead to more sustainable use of natural resources. The CAC treatment also performed well in reducing forest use, while the PES treatment had limited effects. These results are, however, preliminary, and more research is needed in other locations and also to test the sensitivity of parameters and treatment design. While the treatments included try to mimic real-world measures, changes in the level of payment (PES), the probability of detection and the level of penalty (CAC) and the use of social sanctions (CFM) may affect the relative outcomes of the three main treatments. Individual characteristics are predictive of behaviour. Female participants harvest significantly more than men, and younger participants harvest more than older participants (up to 42 years). Gender also differentiates the impacts of the treatments. These findings are somewhat contradicting earlier theory (especially on women and natural resources), but is consistent with results in some earlier experiments. The composition of local forest users, and of the forest user groups or similar committees commonly established to manage forests, will therefore influence the effectiveness of conservation measures.

We decompose external validity into behavioural and treatment validity. The strong and significant correlation between behaviour in the experiment and in the stated real life forest use lends support to the behavioural validity of the experiment, while the experiment's treatment validity is harder to test. Even though the treatments appear to incorporate the core idea of three forest conservation measures, there might be non-relevant factors affecting the treatment impacts, and the impacts might be parameter sensitive.

Forest specific FFEs is a useful method for evaluating the impact of forest conservation measures on the behaviour of local forest users. With higher external validity than lab experiments and less costly (morally, politically, and financially) to implement than RCTs, the method can be viewed as a meeting point between the two. FFEs should become an important complement to traditional impact assessment, in particular as an ex-ante evaluation of potential conservation measures.

Acknowledgements
We appreciate the financial support from the Climate Change Impacts, Adaptation and Mitigation (CCIAM) programme, funded by Norad. We are also grateful for comments to draft versions of the paper by Frode Alfnes, Maren Elise Bachke, Caroline Wang Gierløff, Eirik Romstad, and for field assistance provided by Donatha Dominic, Lucas Kwiyega, John

F. Kessy, and Yonika M. Ngaga at Sokoine University of Agriculture in Tanzania.

References
Agarwal, B. (2010). *Gender and green governance: the political economy of women's presence within and beyond community forest*. New Delhi: Oxford University Press.

Angelsen, A. (Ed.). (2009). *Realising REDD+: National strategy and policy options*. Bogor: CIFOR.

Angelsen, A. (2010). Policies for reduced deforestation and their impact on agricultural production. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(46), 19639–19644. doi:10.1073/pnas.0912014107

Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ*, *340*(c2220). doi:10.1136/bmj.c2220

Bargh, J. A. (2006). What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology*, *36*(2), 147–168. doi:10.1002/ejsp.336

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*(2), 230–244. doi:10.1037/0022-3514.71.2.230

Blomley, T., & Iddi, S. (2009). *Participatory Forest Management in Tanzania: 1993-2009*. Dar es Salaam: Ministry of Natural Resources and Tourism, Forestry and Beekeeping Division.

Blomley, T., Lukumbuzya, T., & Brodning, G. (2011). *Participatory forest management and REDD+ in Tanzania*. Washington D.C.: World Bank.

Bowles, S. (2008). Policies designed for self-interested citizens may undermine "the moral sentiments": evidence from economic experiments. *Science*, *320*(5883), 1605–1609. doi:10.1126/science.1152110

Brockington, D. (2007). Forests, community conservation, and local government performance: The village forest reserves of Tanzania. *Society & Natural Resources*, *20*(9), 835–848. doi:10.1080/08941920701460366

Brown-Kruse, J., & Hummels, D. (1993). Gender effects in laboratory public goods contribution. *Journal of Economic Behavior & Organization*, *22*(3), 255–267. doi:10.1016/0167-2681(93)90001-6

Camerer, C. F. (2011). The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List. *Working Paper*. doi:10.2139/ssrn.1977749

Cardenas, J.-C. (2000). How do groups solve local commons dilemmas? Lessons from experimental economics in the field. *Environment, Development and Sustainability*, *2*(3-4), 305–322. doi:10.1023/A:1011422313042

Cardenas, J.-C. (2004). Norms from outside and from inside: an experimental analysis on the governance of local ecosystems. *Forest Policy and Economics*, *6*(3-4), 229–241. doi:10.1016/j.forpol.2004.03.006

Cardenas, J.-C., Dreber, A., von Essen, E., & Ranehill, E. (2011). Gender and cooperation in children: Experiments in Colombia and Sweden. *SSE/EFI Working Paper Series in Economics and Finance*, *No.735*.

Carpenter, J., Connolly, C., & Myers, C. K. (2008). Altruistic behavior in a representative dictator experiment. *Experimental Economics*, *11*(3), 282–298. doi:10.1007/s10683-007-9193-x

Cilliers, J., Dube, O., & Siddiqi, B. (2012, November). "White man`s burden"? A field experiment on generosity and foreigner presence. *Paper Presented at the Berkeley Symposium on Economic Experiments in Developing Countries (SEEDEC)*.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, *48*(2), 424–455. doi:10.1257/jel.48.2.424

Duflo, E., Glennerster, R., & Kremer, M. (2008). Using randomization in development economics research: A toolkit. In T. P. Schultz, & J. A. Strauss (Eds.), *Handbook of development economics* (Vol. 4.). Amsterdam: North-Holland.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, *326*(5952), 535–538. doi:10.1126/science.1168244

FAO. (2011). *State of the world's forests*. Rome: The Food and Agriculture Organization (FAO).

FAO. (2012). *State of the world's forests*. Rome: The Food and Agriculture Organization (FAO).

Ferraro, P. J., & Price, M. K. (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *Review of Economics and Statistics*, *95*(1), 64–73.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics*, *115*(3), 791–
810. doi:10.1162/003355300554917

Grossmann, I., Na, J., Varnum, M. E. W., Park, D. C., Kitayama, S., & Nisbett, R. E. (2010). Reasoning about social conflicts improves into old age. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(16), 7246–7250. doi:10.1073/pnas.1001715107

Harrison, G. W., Lau, M. I., & Elisabet Rutström, E. (2009). Risk attitudes, randomization to treatment, and self-selection into experiments. *Journal of Economic Behavior & Organization*, *70*(3), 498–507. doi:10.1016/j.jebo.2008.02.011

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009– 1055. doi:10.1257/0022051043004577

Harrison, G. W., List, J. A., & Towe, C. (2007). Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica*, *75*(2), 433–
458. doi:10.1111/j.1468-0262.2006.00753.x

Heckman, J. J., & Smith, J. A. (1995). Assessing the case for social experiments. *Journal of Economic Perspectives*, *9*(2), 85–110. doi:10.1257/jep.9.2.85

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2-3), 61–83; discussion 83–135. doi:10.1017/S0140525X0999152X

IPCC. (2013). Summary for policymakers. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge and New York: Cambridge University Press.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*(2), 153–174. doi:10.1257/jep.21.2.153

Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioural economics. *The Economic Journal*, *109*(453), 25–34. doi:10.1111/1468-0297.00400

Lusk, J. L., Pruitt, J. R., & Norwood, B. (2006). External validity of a framed field experiment. *Economics Letters*, *93*(2), 285–290. doi:10.1016/j.econlet.2006.05.016

Meier, S., & Frey, B. S. (2004). Matching donations: Subsidizing charitable giving in a field experiment. *Zurich IEER Working Paper*, *No.181*.

Miguel, E., & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217. doi:10.1111/j.1468-0262.2004.00481.x

Muradian, R., Arsel, M., Pellegrini, L., Adaman, F., Aguilar, B., Agarwal, B., … Urama, K. (2013). Payments for ecosystem services and the fatal attraction of win-win solutions. *Conservation Letters*, *6*(4), 274–279. doi:10.1111/j.1755-263X.2012.00309.x

Mwakalobo, A. B. S., Kajembe, G. C., Silayo, D. S., Nzunda, E., Zahabu, E., Maliondo, S., & Kimaro, D. N. (2011). *REDD and sustainable development – perspective from Tanzania*. London: International Institute for Environment and Development (IIED).

Ostrom, E. (2006). The value-added of laboratory experiments for the study of institutions and common-pool resources. *Journal of Economic Behavior & Organization*, *61*(2), 149–163. doi:10.1016/j.jebo.2005.02.008

Ostrom, E., Gardner, R., & Walker, J. M. (1994). *Rules, games, and common-pool resources*. Ann Arbor: University of Michigan Press.

Pritchett, L. (2008). It pays to be ignorant: A simple political economy of rigorous program evaluation. In W. Easterly (Ed.), *Reinventing foreign aid*. Cambridge and London: MIT Press.

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

Rodrik, D., & Rosenzweig, M. R. (2010). Preface: Development policy and development economics: An introduction. In D. Rodrik, & M. R. Rosenzweig (Eds.), *Handbook of development economics* (Vol. 5). Amsterdam: North-Holland.

Schwieren, C., & Sutter, M. (2008). Trust in cooperation or ability? An experimental study on gender differences. *Economics Letters*, *99*(3), 494–497. doi:10.1016/j.econlet.2007.09.033

Shiva, V. (1989). *Staying alive: Women, ecology, and development*. London: Zed books.

Stern, N. (2006). *Stern Review: The economics of climate change*. Cambridge: Cambridge University Press.

TNRF. (2009). *Using the nation's resources to reduce poverty?* Arusha: Tanzania Natural Resource Forum (TNRF).

UNFCCC. (2007). *Decision 2/CP.13: Reducing emissions from deforestation in developing countries: approaches to stimulate action*. Bonn: United Nations Framework Convention on Climate Change (UNFCCC).

URT. (2012). *National strategy for reduced emissions from deforestation and forest degradation (REDD+)*. Dar es Salaam: Vice president's office, United Republic of Tanzania (URT).

Vohs, K. D., Mead, N. L., & Goode, M. R. (2006). The psychological consequences of money. *Science*, *314*(5802), 1154–1156. doi:10.1126/science.1132491

Voors, M., Turley, T., Kontoleon, A., Bulte, E., & List, J. A. (2012). Exploring whether behavior in context-free experiments is predictive of behavior in the field: Evidence from lab and field experiments in rural Sierra Leone. *Economics Letters*, *114*(3), 308–311. doi:10.1016/j.econlet.2011.10.016

World Bank. (2008). *Putting Tanzania's hidden economy to work: Reform, management, and protection of its natural resource sector*. Washington D.C.: World Bank. doi:10.1596/978-0-8213-7462-7