Strategic Behavior and the Environment, 2017, 7: 179–219

To Cheat or Not? Results from Behavioral Experiments on Self-monitoring in Vietnam

Rohit Jindal[1], Joe Arvai[2], Delia Catacutan[3] and Dam Viet Bac[4]

[1] *Department of Decision Sciences, School of Business, MacEwan University, Edmonton, Canada; JindalR@macewan.ca*
[2] *University of Michigan, Ann Arbor, MI, USA*
[3] *Decision Research, Eugene, OR, USA*
[4] *World Agroforestry Centre, Hanoi, Vietnam*

ABSTRACT

The central question that we ask in this paper is: do people cheat or behave dishonestly when they are unsupervised? This question is motivated by our work on environmental management programs in developing countries where local conservation effort can be either too expensive or difficult to monitor by outsiders. We combine responses from a household survey with results from a set of field experiments in rural Vietnam to examine how people behave when they are unsupervised, and how well are they able to predict the behavior of others in their community. Both the survey and the field experiments are structured along the lines of an environmental project in which local people receive an incentive for providing their time and labor to undertake project activities. Our total sample size is 400 with experimental treatments varying by group size, whether people receive cash or in-kind incentives, and whether they are monitored. In contrast to existing studies that predict compulsory cheating, we find very little cheating behavior among our subjects. Moreover, people can accurately predict the behavior of others except in a scenario when people collect

ISSN 1944-012X; DOI 10.1561/102.00000080
© 2017 R. Jindal, J. Arvai, D. Catacutan and D. V. Bac

cash incentives for others. In that case, people become extra cautious than what they were predicted to be. We explain these results by invocating the theory of self-concept maintenance and conclude with a discussion on potential applications of these results.

Keywords: PES; self-monitoring; Vietnam; experiments; cheating

JEL Codes: C91, C93, Q56, Q57

1 Introduction

Strategic behavior in the absence of external supervision has been an important area of research for social and behavioral scientists due to its ubiquity in several walks of life such as in public policy, business management, accounting, and self-reporting to government agencies (Ariely, 2012; Bazerman *et al.*, 2002; Fisman and Miguel, 2007; Polinsky and Shavell, 1984). One issue at stake is deliberate cheating or opportunist behavior when people are not being observed by others (Mazar *et al.*, 2008). Another is conflict of interest when people cheat almost unconsciously because they may benefit from a certain behavior or by taking a particular position (Moore *et al.*, 2006). They ramp up information that supports their internal rationalization while discounting contradictory evidence. Presence or absence of supervision makes little difference because people are convinced that they are completely truthful in their behavior (Bazerman *et al.*, 2002). This explains why doctors may recommend unnecessary and costly procedures on patients which they themselves are likely to perform. Similarly, accounting firms may become complacent when auditing their own clients (Moore *et al.*, 2006).

Both these kinds of strategic and dishonest behaviors are very relevant for environmental policy. There are several environmental programs in which economic incentives for program participants are directly dependent on adoption of conservation activities at the local level, but that are difficult to monitor by external agents. In addition, group settings in which individuals have an incentive to partake disproportionate benefits can threaten the success of the program for the entire

group. For example, under the program on Reducing Emissions from Deforestation and Forest Degradation (REDD), local communities in developing countries receive external payments for local conservation efforts that are not easy to monitor (Danielsen *et al.*, 2011; Kenney *et al.*, 2015). This gives an incentive to community members to collect more benefits than their fair share, jeopardizing the success of the entire program.

Self-monitoring is also crucial for environmental programs in which local land stewards receive payments for producing valuable ecosystem services such as carbon sequestration and erosion control. Known as payments for ecosystem services (PES), such payment schemes have become common in many parts of the world and are the focus of the Paris Climate Agreement under its performance-based emission reduction system (Jindal *et al.*, 2007; Wong *et al.*, 2016). It is expected that when fully operational, such PES-type emission reduction activities will be worth billions of dollars, with most of the money flowing to developing countries that contain large chunks of tropical forests. Even though most of these projects will be monitored through state of the art remote sensing tools (Bucki *et al.*, 2012), not all local activities could be readily verifiable by outsiders. It is possible that program participants may share external payments unequally, or some land owners may exploit disproportionate share of local resources when no one is watching. Such actions will threaten the long-term sustainability of environmental projects and may even limit donors' ability to support them. In recent years, the environmental management literature has begun to examine the possibility of misreporting by local stakeholders in environment and development projects (Holck, 2008). The focus within this literature has primarily been on lack of technical expertise within local communities, which could potentially be addressed by training people to take more accurate measurements, or by adopting more stringent protocols (Danielsen *et al.*, 2011; Skutsch, 2012). However, the behavioral aspects of self-monitoring including the possibility of cheating by program participants have remained unexplored so far.

We attempt to address this gap in literature by reporting behavioral outcomes from a set of field experiments in which we follow a typical PES scenario whereby people receive non-trivial incentives for carrying out a specific activity. Although it would have been ideal to design experiments around actual PES schemes, we were unable to do this

due to resource constraints. As a close substitute, we carried out a set of field experiments using PES-like scenarios (by varying the level of monitoring, kind of incentive, and size of the subject group), that help us to observe people's behavior when they have an incentive to cheat.

Following the tradition of behavioral economists and psychologists, we compare the results of these experiments with responses from a survey questionnaire in which respondents are asked to predict the behavior of others under similar scenarios (Epley and Dunning, 2000). The combination of field experiments and survey questionnaire helps us in generalizing our results while extending some of the previous experiments from lab settings to the field setting in a developing country context.

Vietnam provides a relevant field context for the study. It is one of the 13 countries where the United Nations' REDD program is being piloted. Besides, the country has launched a major PES initiative under which local communities will receive payments for conserving large water basins. There is a strong push among program implementers to look at the viability of self-monitoring by local communities in these programs (Hoang *et al.*, 2013). Vietnam has also been the focus of some recent studies on trust and cooperation among community members (Bouma *et al.*, 2008; Nguyen *et al.*, 2012; Parks and Vu, 1994). While Nguyen *et al.* (2012) find that trust level among their Vietnamese subjects is similar to the level reported for other countries, Parks and Vu (1994) report a much higher level of cooperation among Vietnamese than amongst western citizens. We do not replicate these studies but some of our research questions are similar in scope which helps us to add more depth to the current discussion on trust and cooperation in social psychology. Overall, we believe that our research is relevant for both policy makers in Vietnam and for a wider audience interested in behavioral outcomes under self-monitoring.

2 Literature Review and Research Hypotheses

One line of research on human behavior has focused on the propensity to cheat among people when they are not being watched. Many researchers suggest that individuals behave dishonestly if benefits of doing so are greater than the expected costs of cheating (the probability of being

caught times the financial penalty for cheating). According to this rational model of human behavior (Becker, 1968; Polinsky and Shavell, 1984), we should expect people to cheat when they have an opportunity to do so. Recent evidence from lab experiments, however, belies this prediction. Many of these experiments show that even though most people do cheat when unsupervised, the extent of their cheating is much less than what is predicted by Becker's model (Ariely, 2012; Mazar *et al.*, 2008). A more comprehensive study conducted with college students across 16 countries finds similar results (Pascual-Ezama *et al.*, 2015). When tempted to claim a chocolate by misreporting the outcome of a coin, most individuals show honest behavior.

This apparent dichotomy is explained by the theory of self-concept maintenance (Mazar *et al.*, 2008). This theory suggests that humans try to limit their cheating behavior by striking a balance between two opposing forces — temptation to gain from cheating versus restraining themselves to uphold their self-image as honest beings. This theory is a modification of the concept of cognitive dissonance in psychology that suggests that individuals try to strike a consistency between their internal beliefs and their external behavior (Ploner and Regner, 2013). When confronted by new information that contradicts their internal beliefs, people change how they perceive the information rather than their belief system. The theory of self-concept maintenance suggests that individuals create a moral wiggle room that lets them maintain their internal belief of being honest by cheating a little bit when they have an opportunity to do so, but not as much as they potentially could (Mazar *et al.*, 2008).

Even though the two lines of work (rational model of cheating versus self-concept maintenance) differ in their line of reasoning, they both predict that when confronted with temptation to cheat, most people will cheat. The difference lies in the level of cheating. These studies suggest that cheating due to moral conflict is independent of the specific context of the subjects (Ariely, 2012). Though a major limitation with generalizing their results is that most of these are based on lab experiments with college students (Mazar *et al.*, 2008; Pascual-Ezama *et al.*, 2015).

A contrasting view on human behavior relates to existence of social norms such as trust and cooperative behavior and how well they are shared among people (Kerr *et al.*, 2012). Conformity to norms is

conditional on people's expectations of what others will do and how they themselves are expected to behave, making social norms local and context dependent (Bicchieri, 2008). If these norms are well internalized, people are less likely to violate them even under pressure (Campbell, 1964). Context thus plays an important role in explaining people's behavior (Levine, 2003). Evidence of this phenomenon comes from several studies: Fisman and Miguel (2007) report higher number of parking violations among United Nations officials who belong to countries with more propensity to be corrupt. Similarly, experiments among urban slums across Bangkok and Ho Chi Minh show significantly higher level of trust and cooperation among Vietnamese subjects, which is explained by existence of a collectivist culture that promotes certain kinds of pro-social norms (Cárdenas and Ostrom, 2004).

Given this background, the questions that arise in context of environmental programs are whether cheating is ubiquitous, and whether the context under which people operate affects their behavior. If cheating is always to be expected, environmental programs in general, and PES schemes in particular, would require strict external monitoring without which the programs should expect significant levels of non-compliance and misreporting by local participants. On the other hand, if human behavior is affected by the context in which it operates, then program managers would need to take local factors, such as social norms into account, especially when external monitoring becomes difficult to enforce.

This brings us to our first research hypothesis:

H1: People will behave dishonestly when they are unsupervised.

In our research, we test this hypothesis through a combination of survey and artefactual experiments in the field. More and more studies in behavioral economics and social psychology tend to follow this method to test subjects' predictions about their behavior against their actual behavior. The combination thus provides more enhanced results (e.g. see Qin *et al.*, 2011). In case of incentive-based environmental management programs such as PES, we believe there are three specific situations that need to be considered: (i) effect of cash versus in-kind incentives on people's propensity to be honest, (ii) effect of individual versus group monitoring on people's behavior, and (iii) effect of self-monitoring versus monitoring others in the group.

(i) Cash versus in-kind incentives

Previous research shows that humans respond differently to cash and in-kind incentives even when the size of the incentive is the same. This is due to a difference in perception whereby people are calculative when monetary incentives are involved while they are likely to be more ethical when responding to non-monetary incentives (Heyman and Ariely, 2004). However, other studies contradict this and suggest that cheating can be more pronounced when in-kind incentives are present (Ariely, 2012; Mazar *et al.*, 2008). This happens because people find it easier to rationalize cheating with in-kind benefits. Cheating with cash is more difficult to rationalize where even a small discrepancy is clearly noticeable. This brings us to our second hypothesis — when external supervision is absent:

H2: People are more dishonest when facing in-kind incentives versus when facing cash incentives.

(ii) Individual versus group monitoring

When individuals form groups, they usually lose some level of control as the process of group dynamics seems to supersede individual behavior (Forsyth, 2009). And when groups have well evolved social norms towards cooperation, individual members continue to cooperate for the overall public good (Ostrom, 2000). There is evidence that this norm-based cooperation can translate to pro-social behavior and good citizenship across other aspects of civic life too (Levine, 2003; Cleaver, 2000; Steblay, 1987). This implies that in the presence of strong social norms towards pro-social behavior, both individuals and groups should be equally honest. However, Wiltermuth (2011) finds that once people cheat, a group can be more dishonest than individuals as group members rationalize their unethical behavior by splitting the spoils of their dishonesty with other members of the group. This leads to our third hypothesis:

H3: People are more dishonest when they are in groups than when they work individually.

However, if there exist strong social norms towards honest behavior, both individuals and groups can be expected to behave similarly.

(iii) Self-versus monitoring others

Motivational biases often make people behave differently when looking at their own gains versus gains and losses for others (Wilson *et al.*, 2008). Due to a difference in perception, people may see their own losses as unique while underestimating the same for others. Dana *et al.* (2007) propose that people in general like to be viewed as fair. However, when there is some flexibility for opportunistic behavior, they are harsher in their judgement of others than of themselves.

Our combination of survey and experiments allows us to compare people's predictions about their behavior with their actions (Mazar *et al.*, 2008). Experimental evidence shows that people predict others' behaviors more accurately compared to predicting their own behavior (Epley and Dunning, 2006). For example, in another study, Epley and Dunning (2000) find that while 83% of the survey respondents said they would cooperate for a public cause, only 43% did so. The same respondents felt that only 56% of their peers would contribute to this cause, which turned out to be relatively close to the actual behavior. The researchers hypothesize that this happens because people often have a more charitable opinion about themselves compared to a more accurate view of others (Epley and Dunning, 2000, 2006). This leads to our final hypothesis:

H4: People are more dishonest when they are monitoring themselves compared to when they are monitoring others.

3 Household Survey: Responses to Incentive Scenarios in Rural Vietnam

The main purpose of the survey was to collect people's predictions regarding their behavior under different scenarios. As we describe below, these scenarios were constructed to mimic a PES project under which people receive economic incentives for taking up conservation that requires time and effort. Responses to these survey questions were compared with outcomes from artefactual experiments with a similar set of respondents (though in different villages). The survey followed an approach that has been successfully tested by other researchers whereby randomly selected respondents are asked to articulate their

preference among a given set of scenarios (e.g. see Kerr *et al.*, 2012; Lusk and Norwood, 2009b). Aggregation of these responses indicates the proportion of community members that prefer a specific outcome from the choices that are given.

The second component of the survey was designed with a purpose to collect information on the existence of pro-social norms of trust and cooperation within the community. Questions that ask respondents the extent to which people trust their community members, and the frequency with which people help others in time of need, provide an estimate of the extent to which trust and cooperation have been internalized by community members (Kerr *et al.*, 2012; Qin *et al.*, 2011). However, a common problem with preference elicitation is social desirability bias; respondents may provide answers that show them in the best possible light on sensitive issues, thus distorting the overall results (Fisher, 1993). To address this bias, Lusk and Norwood (2009a,b) propose an inferred valuation approach which involves asking participants to articulate preferences of their neighbors or community members rather than their own opinions. For example, words such as "I will . . ." are replaced by, "My neighbor will . . .". Because the questions are indirect or inferred, there is less motivation for the respondent to look good to the interviewer (Lusk and Norwood, 2009b). This method is based on an indirect line of questioning tested by Fisher (1993) and has been found to be effective in removing social desirability bias under certain conditions. Epley and Dunning's (2000; 2006) work also shows that people are more accurate in predicting the behavior of others than their own, so an inferred or indirect form of questioning may be more effective in getting correct responses.

3.1 Research Site

The survey was conducted in selected villages in Bac Kan province in Vietnam. Bac Kan is a mountainous province, located 170 km north of the capital city of Hanoi and a key site for the national PES program. Average household income is low in rural Bac Kan, where the main source of income is agriculture. However, with more than 70% of the land area covered by forests, national policy makers foresee potential to improve local incomes through sustainable management of forests, and have shortlisted Bac Kan as a candidate for forest-based carbon management program promoted by the United Nations (Hoang *et al.*, 2013).

Table 1: Details of respondents for survey questionnaire.

Village	Number of respondents	Average age (years)
Ban Cuon	29	43.6
Ban Ken	42	40.3
Na Muc	11	39.5
Na Ngoa	45	37.2
Phieng Bang	30	41.2
Total	**157**	**40.1 (average)**

3.2 Research Methods

Survey data were collected from 157 respondents across five villages of Na Ngoa, Phieng Bang, Ban Ken, Na Muc, and Ban Cuon (Table 1). The average age of these respondents was 40.1 years. 38.8 percent of the respondents were female. Since the survey respondents represented similar proportion of the total population, smaller villages such as Na Muc and Ban Cuon provided a lower share of the overall sample. The composition of respondents was similar across the five villages (one-way ANOVA for difference in age of respondents across villages was not significant; $F-\text{stat} = 1.7$; $p > 0.15$).

Household contact was established through village leaders. The questionnaire was administered to the household head or to the eldest household member (in the absence of the household head). The questionnaire consisted of three main sections: the first on demographic information, the second on the level of trust within community members. For this, participants were presented with two multiple-choice questions that asked them to rate how much did people in their village trust fellow villagers by selecting one of the four possible responses — no trust, only sometimes, everyone trusts, and don't know (adapted from Kerr *et al.*, 2012). The second question was how often did people help others in time of need? Possible responses were — never, sometimes, often, and every time. Section three of the questionnaire asked participants to predict outcomes under a hypothetical set of scenarios which included supervised versus unsupervised situations. To address any potential social desirability bias, instead of asking participants about their own

behavior (Lusk and Norwood, 2009a), the questions asked participants to predict the behavior of their neighbor under these scenarios. In each of the scenarios, participants were told to suppose people in their village were asked to fill in a questionnaire after which they could receive cash or an in-kind incentive. In order to compensate for the time taken to fill the questionnaire, these people would receive VND 70,000 (∼US 3.5) per person (equivalent to one day's wage rate). Alternatively, the in-kind incentive that was offered was 750 grams of maize seeds per person, the monetary value of which was VND 70,000. These incentives and the related scenarios mimicked a PES-like program whereby farmers provide time and labor to plant trees (or protect forests) in return for adequate compensation.

There were four sets of questions in the survey, each covering four scenarios for a total of 16 scenarios (Table 2). These same scenarios were also covered as treatments during the field experiments. In the survey, each participant had to respond to one of the randomly selected scenarios. Each set of scenarios started with informing participants to suppose their neighbor could receive an incentive for filling up a survey questionnaire. In set I's first scenario (scenario A), respondents were asked how much maize seeds would their neighbor collect when they were asked to go alone into a room and collect 750 grams. The four possible choices were — (i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams. In set II's second scenario (scenario B), the respondent was asked to predict the same neighbor's behavior if they also had to collect seeds for a resident in another village. The choices were the same as before — (i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams. In set I's third scenario (scenario C), the respondent predicted how much the neighbor would collect if instead of seeds, they had to collect VND 70,000, of cash — (i) less than VND 50,000, (ii) VND 51,000–69,000, (iii) VND 70,000, (iv) VND 71,000 — 90,000, and (v) more than VND 90,000. In set I's fourth scenario (scenario D), the respondent had to predict how much cash would their neighbor collect for a resident in a neighboring village — (i) less than VND 50,000, (ii) VND 51,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000.

As Table 2 shows, the four sets (sets I–IV) varied by the type of incentive available to the neighbor (seeds, cash), and whether the

Table 2: Details of the four scenarios presented to respondents in the survey questionnaire.

	Set I	Set II	Set III	Set IV
Scenario A	How much seeds would your neighbor collect for herself?	How much seeds would your neighbor collect for herself?	How much cash would your neighbor collect for herself?	How much seeds would the two neighbors collect for themselves?
	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	(i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams
Scenario B	How much seeds would your neighbor collect for someone else?	How much seeds would your neighbor collect for someone else?	How much cash would your neighbor collect for someone else?	How much seeds would the two neighbors collect for someone else?
	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	(i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams
Scenario C	How much cash would your neighbor collect for herself?	How much seeds would the two neighbors collect for themselves?	How much cash would the two neighbors collect for themselves?	How much cash would the two neighbors collect for themselves?
	(i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	(i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	(i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	(i) less than VND 50,000 (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000

(*Continued*)

Table 2: (*Continued*)

	Set I	Set II	Set III	Set IV
Scenario D	How much cash would your neighbor collect for someone else? (i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	How much seeds would the two neighbors collect for someone else? (i) less than 600 grams, (ii) 600–749 grams, (iii) 750 grams, (iv) 751–900 grams, and (v) more than 900 grams	How much cash would the two neighbors collect for someone else? (i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000, (iv) VND 71,000–90,000, and (v) more than VND 90,000	How much cash would the two neighbors collect for someone else? (i) less than VND 50,000, (ii) VND 50,000–69,000, (iii) VND 70,000 (iv) VND 71,000–90,000, and (v) more than VND 90,000

neighbor went alone or with someone else as part of a pair. The questionnaire followed the same design as used in the behavioral experiments (Tables 3 and 4). Apart from recording behavioral outcomes for each scenario, the respondents were also asked to state the reason behind their response to the scenario questions, the three choices being — (i) because the neighbor was honest, (ii) because the neighbor would follow instructions, and (iii) because it was okay to pick a different level of incentive.

3.3 Results

Each participant responded to one randomly selected set covering four different scenarios — 157 respondents resulting in 628 observations. Eighty-four percent of respondents said that people in their village trusted everyone, while 11 percent said that they trusted others sometimes (Figure 1a). Only 4.5 percent of the respondents said they did not know the extent of trust in their village while less than 1 percent thought there was no trust among fellow villagers. Similarly, 64 percent of respondents said people often helped others, while 7 percent said that people helped others every time. Twenty-eight percent said that people helped each other sometimes, while only 1 percent of the respondents said that people never helped others (Figure 1b). These responses

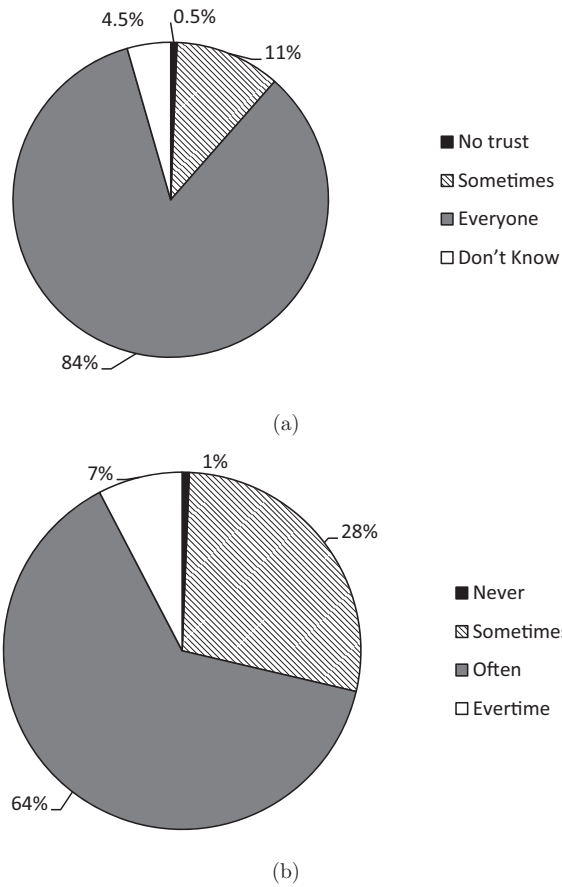

Figure 1: Survey responses regarding level of trust and helping behavior in the community. (a) Level of trust (b) Frequency of helping behavior.

indicate that there exists a very high level of trust in the area along with a strong social norm of helping others in time of need.

In response to the four scenario questions, 99 percent of the respondents said that their neighbor would collect the correct level of incentive even when not monitored. This was an extremely strong result and significantly different from the behavioral outcomes observed in many of the previous studies on cheating behavior (Ariely, 2012; Fisman and Miguel, 2007; Mazar *et al.*, 2008). When asked the reason for this expected behavior, 67 percent of the respondents said that this was

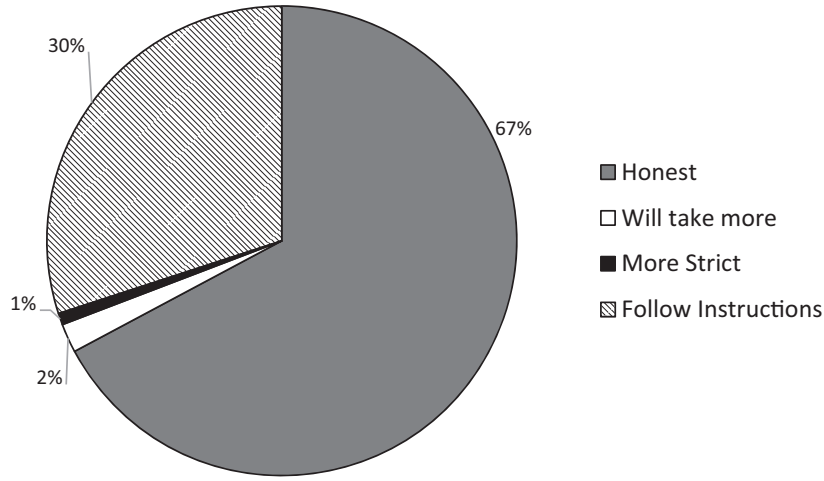

Figure 2: Predictions regarding behavior in survey responses.

because they expected their neighbor to be honest, while 30 percent felt that this was because their neighbor would follow the instructions that mentioned clearly the amount of incentive to be collected (Table 2). Interestingly, 2 percent of respondents felt that their neighbors could collect more than asked, while about 1 percent said that they expected their neighbor to be stricter when collecting incentive for others.

3.4 Discussion

Taken together, these results indicate that there exists high level of trust and honesty in our study area. We discuss these results below before moving on to evidence we collected for hypothesis testing.

The trust level measured in our study is much higher than the one estimated through the trust game reported in Nguyen *et al.* (2012). In their experiment, trust is measured by the proportion of the amount sent by one player to another. In line with some other studies that use this game, these researchers found that on average, people sent only 49.25 percent of their initial endowment to the other person (Nguyen *et al.*, 2012). However, a more comprehensive review of related studies shows considerable variation in measurement of trust even for Vietnamese subjects. For example, Parks and Vu (1994) found that people

of Vietnamese origin were substantially more cooperative than their American counterparts, and that their rate of cooperation did not decline with repetition. Similarly, in a study conducted across urban slums in Thailand and Vietnam, Vietnamese subjects were found to be significantly more trusting and cooperative (Carpenter *et al.*, 2004). On average, Vietnamese subjects contributed 7.41 out of their endowment of 10 bills, which was significantly higher than Thai subjects (6.72 out of 10). The researchers conclude that not only were contributions higher among Vietnamese subjects, but also the variance of their contributions was lower indicating players in Vietnam conform much more than in Thailand (Carpenter *et al.*, 2004). Given the close-knit society in our study area and the existence of a collectivist culture, we believe that the high level of trust we find is not erroneous, and is an appropriate representation of the helping and trusting nature of the people in the community.

Similarly, the strong honest behavior predicted by our respondent group is not an extreme outlier. In their experiment across 16 countries, Pascual-Ezama *et al.* (2015) found relatively high levels of honesty. In our study, 67 percent of the respondents said that their neighbor would collect the correct level of incentive because they were honest, and another 30 percent felt that their neighbor would follow instructions. While the former reason indicates the presence of a social norm of being honest, the latter points out a rule abiding society with a history of collectivist culture (Carpenter *et al.*, 2004). If this high level of honesty were to be corroborated by the field experiments, it would indicate the importance of local context in shaping human behavior.

A potential limitation of these results is that our respondent group could be following another social norm of not saying anything bad about their neighbor, and hence the strong positive result in terms of trust and honesty. We return to this issue in Section 5.

4 Field Experiment: Incentive Types and Extent of Supervision in Rural Vietnam

There were two components of our study, the household survey and the field experiments. Combining these two methods enabled us to collect more comprehensive data (Mazar *et al.*, 2008; Qin *et al.*, 2011; Scarpa *et al.*, 2003). For consistent test of research hypotheses, we

Table 3: Details of subjects for field experiment.

Village	Number of respondents	Average age (years)
Choi Moi	52	38.5
Na Deng	49	41.8
Na Diec	50	38.8
Na Du	58	41.7
Pac Ban	34	39.1
Total	**243**	**40.1 (average)**

conducted our experiments using similar scenarios as used in the survey data in Section 3 above. In recent years, field experiments have become increasingly important in environmental economics (Brent *et al.*, 2017; Price, 2014), including their use to evaluate environmental policy decisions in developing countries (Cárdenas and Ostrom, 2004; Kerr *et al.*, 2012; Lopez *et al.*, 2012). However, to the best of our knowledge, the work on cheating behavior has been mainly confined to lab experiments in developed countries (e.g. Mazar *et al.*, 2008; Pascual-Ezama *et al.*, 2015), making our experiments in Vietnam a novel addition to the experimental literature.

A total of 243 people participated in the experiments split across five villages of Choi Moi, Na Deng, Na Diec, Na Du, and Pac Ban located in Na Ri district of Bac Kan province (Table 3). These villages were different from the ones that participated in the household survey. Since the overall study area was quite similar, a comparison of survey and experimental results is possible. The average age of our experimental subjects was 40.1 years, with 43.1 percent of the respondents being female, which is close to the demographic characteristics of our survey respondents. The distribution in age of subjects across the five villages was similar (one-way ANOVA for differences in age was not significant with $F-\text{stat} = 0.60$ and $p > 0.66$).

4.1 Methods

In the field experiment, we selected the same PES-like scenarios that were used in the survey questionnaire. As shown in Table 4, the experiment followed a $2 \times 2 \times 2$ factorial design. The first factor was

Table 4: Design of treatments and sample size for field experiment.

Factor 1: Extent of Monitoring	Factor 2: Kind of incentive/ compensation	Factor 3: Number of people collecting incentive	Total number of participants
Control — Strict Monitoring	Cash	1	14
		2	14
	In-kind (seeds)	1	16
		2	14
No monitoring — Self monitoring by subjects	Cash	1	24
		2	72
	In-kind (seeds)	1	25
		2	64
TOTAL			**243**

Note: All subjects collected incentive for themselves as well as for a resident of a neighboring village.

whether or not participants were supervised by an external monitor when they collected their incentive (Table 4). The second factor was the kind of incentive that was offered to people (cash or 'in-kind'), and the third factor was the number of people collecting the incentive — single or in-pair. In each case, in addition to collecting incentive for themselves, participants also collected incentive for a resident of a neighboring village (monitoring others condition).

Detailed procedures for experimental treatments (non-monitoring condition)

In each of the five villages where the experiments were conducted, residents were invited to a common meeting place to fill up a questionnaire that included questions on local agriculture and forestry practices. Once assembled, the participants were randomly divided into two groups; one group assigned to the 'cash' treatment whereby all members of the group received money as a compensation or incentive for filling up the questionnaire. The other group was assigned to the 'in-kind' treatment whereby the members of this group received hybrid maize seeds as compensation. We selected a variety called NK67, as the in-kind incentive. This variety is quite popular locally. The market value of the seeds was the same as the level of cash incentive. Each of the two groups was

further divided into two sub-groups: in one group, members received individual incentive, and in the second sub-group the members received their incentive in pair. In all, four groups were formed in each village.

In each group, after participants had handed in the completed questionnaire at research site A, they were directed one at a time to a separate research site B to collect the incentive (Figure 3). This participant was instructed that along with his/her own incentive he/she should also collect an equal amount for a resident of a neighboring village. For this, the participant was provided with two empty envelopes — one for collecting his/her own incentive and the other for the neighbor (two plastic bags in case the incentive was in-kind, i.e. seeds), — to be left in a box outside site B from where the external team would collect it later. The participant was informed that the reason for this was that the external team would visit the neighboring village on the following day to conduct a similar questionnaire and the participants in these villages will receive these envelopes (or bags in case of in-kind incentive) as compensation for their time.

In case of cash incentive offered individually, each participant was asked to collect VND 70,000 (∼US 3.5) for himself and the same amount for the neighbor. However, upon arrival, the participant would find a total of VND 250,000 at site B in denominations of VND 10,000. When cash incentive was offered in pairs, two random participants were sent jointly, again to collect VND 70,000 for each of them and another VND 70,000 for the neighbor. In this case, on their arrival, they would find VND 450,000 in the room, in denominations of VND 10,000.

In case of in-kind incentive, the procedures were the same except that cash was replaced by maize seeds. When seeds were offered individually, participants were asked to weigh 750 grams of seeds for themselves and another 750 grams for the neighbor from a seed bag kept at site B. 750 grams were enough to plant $2/3^{rd}$ of a typical maize plot in the area, with their monetary value being equal to VND 70,000. On arriving at site B, the participant would find 3 kg of seeds from which they would weigh 750 grams using a simple weighing scale that was commonly used in the area. For in-kind incentive offered in pairs, the total quantity of seeds kept at site B was 4.5 kg out of which each of the participants could take 750 grams while keeping a similar amount in the box for the neighbor.

In the no-monitoring condition, when participants went to collect their incentive at site B, there was no one to monitor their behavior.

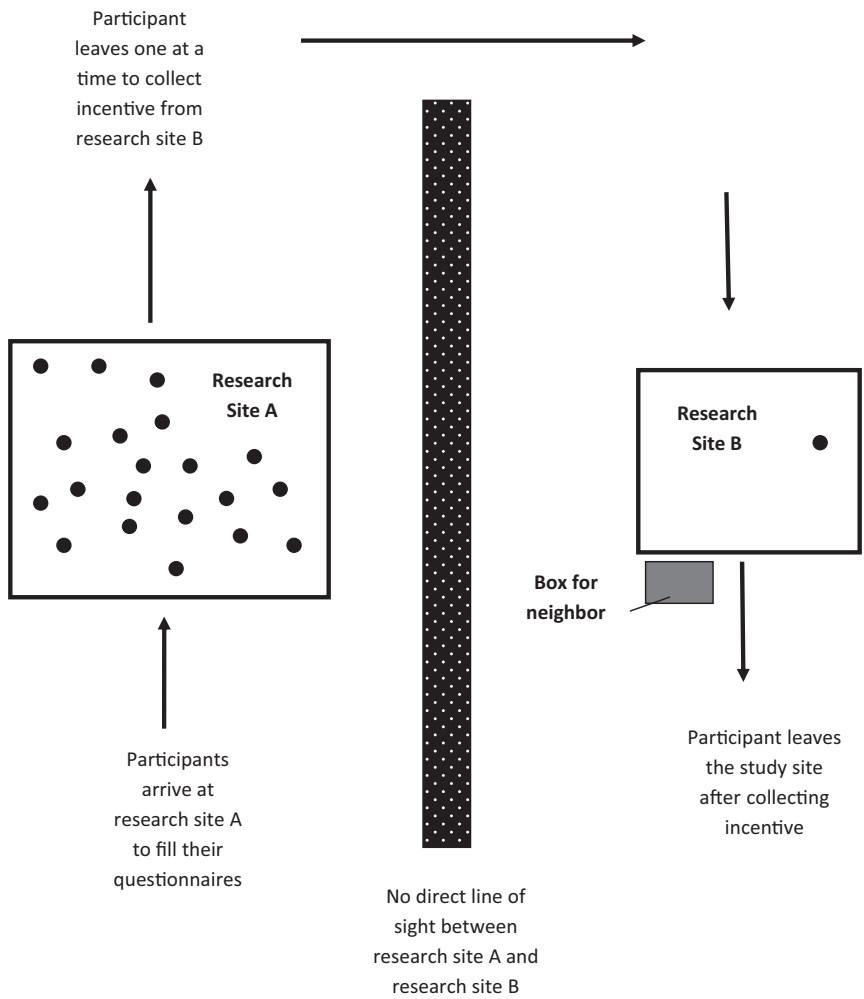

Figure 3: Schematic diagram of the research setting for the field experiment.

Therefore, the two research sites, one where the participants assembled to fill in the questionnaire and the other where they went to pick up their incentive, were chosen with great care by the research team. These two sites were separated from each other by some distance. This meant that participants had no idea of what was happening at site B where the

incentive (money/seeds) was kept. As and when participants completed their questionnaire at site A, they were sent to site B to collect their incentive after which they were free to leave for their house without reporting back to the research team. In addition, when participants arrived on site B to pick up their incentive, they had no expectation about how much cash/seeds were kept there. They had clear instructions to collect their incentive and leave the research site. In the case of in-kind incentives, to ensure that participants' collections remained unseen, the participants were provided with black plastic bags. The time that it took for participants to walk from site A to site B was sufficient for a member of the research team to quickly collect the balance remaining at site B and replace it with a fresh bundle of money or bags of seeds (that were pre-assembled by the research team prior to starting the experiment). This member of the research team remained completely unobserved during this entire process. The actual incentive taken by each participant from site B was estimated by the research team at the end of each day.

In case of cash incentive offered individually, from a total of VND 250,000 an honest participant would collect VND 70,000 for himself and deposit another VND 70,000 in the box outside site B, leaving VND 110,000 behind. In case of a pair of participants collecting cash incentive, from a total of VND 450,000, an honest pair would collect VND 140,000 for themselves (VND 70,000 each) and deposit another VND 70,000 in the box, leaving VND 240,000 behind.

When in-kind incentive was offered individually, from a total of 3 kg, an honest participant would collect 750 grams of seed for himself, deposit 750 grams in the box, thus leaving 1.5 kg behind. In case of pairs, from a total of 4.5 kg, an honest pair would collect 1.5 kg for themselves (750 grams each), deposit 750 grams in the box, while leaving 2.5 kg behind. Any deviations from these numbers were easy to estimate by later measuring how much cash or seeds had been left behind.

Detailed procedures for the control condition (monitoring was present)

Our control consisted of the above four treatments, but in presence of external monitoring. We had the same $2 \times 2 \times 2$ design whereby participants were assigned to cash or in-kind incentive, and went alone

or in pairs to collect their own incentive as well as incentive for a resident in a neighboring village. However, this time, whenever a participant came out of site B, a member of the research team was present to record the amount of money (or seeds) taken by the participant. This was clearly observable to participants as they assembled to fill their questionnaires at research site A. These treatments constituted the strict monitoring case which helped to estimate the level of incentive collected under the presence of external monitoring.

4.2 Results

In our experiments, there were two observations per participant — amount collected for the self and for the neighbor. The same applied when participants went in pairs, amount collected by the pair and the amount left for the neighbor. To analyze participants' behavior under different treatments, we standardized the outcome variable (the amount of incentive collected by a participant) by calculating percent deviation in each case. If a participant collected more than the recommended value, we recorded it as 'over-reporting'. And in case, it the amount collected was lower than recommended, it was recorded as 'under-reporting'. For example, if a respondent collected 1,000 grams of seeds (instead of the recommended 750 grams), it was over-reporting by 33.3 percent $[(1,000 - 750)^*100/750]$. For 500 grams, this would be under-reporting by 33.3 percent. Similarly, for cash incentive, a collection of VND100,000 (instead of VND 70,000) was recorded as over-reporting by 42.9 percent $[(100,000 - 70,000)^*100/70,000]$. In all, there were 324 observations for a total of 243 participants across all treatments.

Result 1: People do not necessarily cheat when unsupervised

How did people behave when they were monitored versus when they were not? Our experimental data shows that when not monitored, excess compensation is collected in 12.3 percent of cases, with an average over-reporting of 17.1 percent for these cases. At a first glance, this seems to support the hypothesis that people cheat when unsupervised. However, we also find that when not monitored, 31.8 percent of the sample collects less compensation than recommended, the average under-reporting for this sub-group being –35.7 percent. Overall, the average compensation

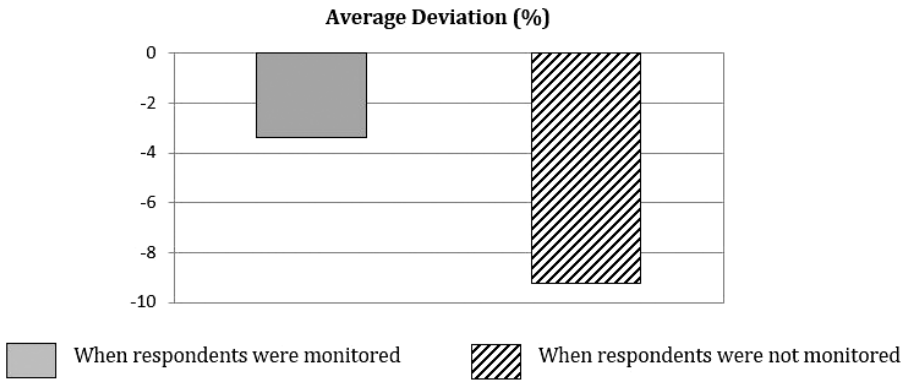

Figure 4: Average deviation in compensation collected (in percentage) across all treatments.

in non-monitoring condition is 5.9 percent lower than under the presence of external monitoring (Figure 4).

Looking closely at the variation in individual responses, even when people are monitored, a small proportion of the participants over-reported, while a larger section under-reported (Figure 5a). A similar trend is observed for participants who are not monitored — a small section over-reported while a much larger number of participants under-reported (Figure 5b). In addition, in the treatment non-monitored, 55.9 percent of the sample collected its designated level of incentive.

We believe that these figures suggest two possible trends: *First*, most respondents are honest in collecting their true level of compensation, and absence of monitoring does not lead to dishonest behavior. The variation in individual responses mostly occurs in the form of weighing error when people collect seeds. This happens because when weighing maize seeds on an analog weighing machine, some people over-weigh and some under-weigh, while a large proportion of the sample weighs close to the correct weight. Therefore, over-reporting of in-kind incentive should be viewed as random error and not as strategic behavior. *Second*, under treatments pertaining to cash incentives, people are stricter in collecting compensation resulting in a much higher magnitude of under-reporting than in over-reporting.

We check for these explanations by analyzing the marginal effects of each treatment through an ordinary least square (OLS) regression

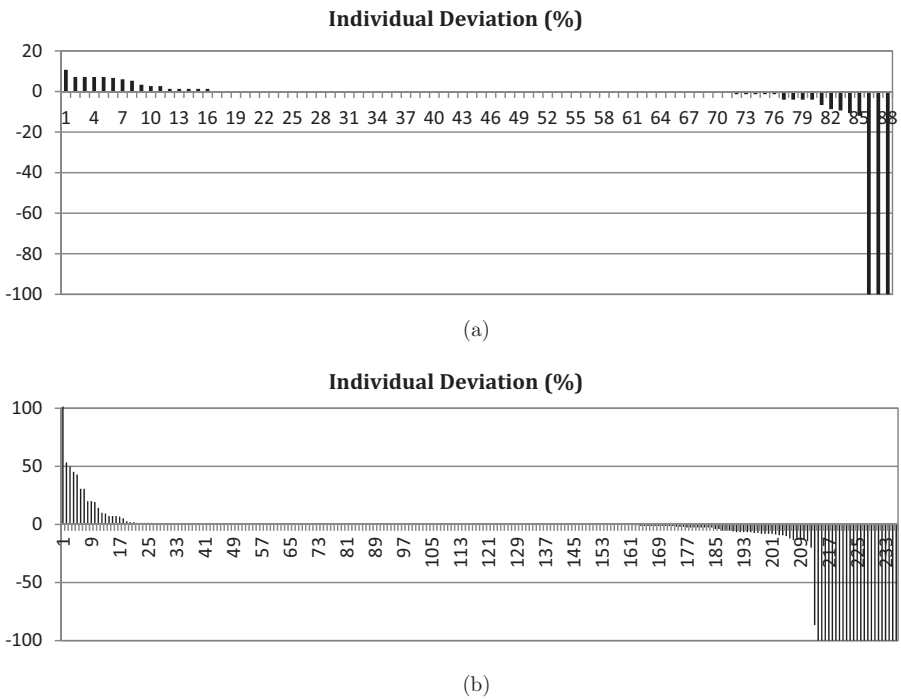

Figure 5: Individual variation in compensation collected. (a) When respondents were monitored (b) When respondents were not monitored.

model. In this model our outcome variable is the percent deviation in compensation collected, while the explanatory variables are the four treatments:

$$Y_i = a + b_1 X_{i1} + b_2 X_{i2} + b_3 X_{i3} + b_4 X_{i4} + e_i \tag{1}$$

where

Y_i = Percent deviation in compensation collected by individual 'i' (>0 under over-reporting and <0 when under-reporting)

X_{i1} = Monitoring treatment (=1 when an individual is monitored and =0 when not monitored)

X_{i2} = Kind of incentive offered (=1 when cash is offered, and =0 when seeds are offered)

X_{i3} = Number of people who collect the compensation (=1 when only one person collects the compensation, =0 when two people go in a pair)

X_{i4} = Whether respondents collect compensation for themselves or for their neighbor (=1 when collecting compensation for themselves, =0 when collecting compensation for the neighbor)

b_1, b_2, b_3, b_4 = Respective coefficients

a = Model constant

e_i = Random error

Table 4 (column 1) presents the results of the regression analysis. It shows that three of the four treatments are significant ($p < 0.10$). When not monitored, people collect 5.9 percent less compensation than when they are monitored (coefficient 5.9, t-stat 1.65). Similarly, people collect less compensation when offered cash versus seeds (coefficient –10.1, t-stat 3.21). The biggest difference is when people collect compensation for themselves versus for others. On average, respondents collect 12.6 percent less compensation for their neighbor than when they collect compensation for themselves (coefficient –12.6, t-stat 4.04). However, there is no difference in outcome when people collect their compensation individually or in pairs (coefficient –1.7, but t-stat 0.52). While we discuss each of these cases in more detail subsequently, we can summarize these outcomes in the form of our first result:

People do not necessarily cheat when they are unsupervised.

Result 2: People are not necessarily more dishonest when operating in groups

Although Wiltermuth (2011) reports that groups tend to cheat more than individuals, we find no such evidence. In our experiment, when

Table 5: Regression results.

Dependent variable: Percentage deviation in incentive collected			
	Variables	OLS with main effects only	OLS with main and selected interaction effects
1.	External monitoring (Yes = 1; No = 0)	5.9* (3.6)	5.9* (3.6)
2.	Cash/In-kind (Cash =1; Seeds = 0)	−10.1*** (3.1)	−1.2 (4.4)
3.	Number of people (1 People = 1; 2 People = 0)	−1.7 (3.2)	−1.7 (3.2)
4.	Self-monitoring/Others (Others = 1; Self = 0)	−12.6*** (3.1)	−3.7 (4.4)
5.	Interaction Self/Others*Cash/In-kind (Others*Cash = 1; Rest = 0)	—	−17.8** (6.2)
	Constant	2.9 (3.2)	−1.6 (3.6)
	Sample size	324	324
	Prob. > F	7.38***	7.70***
	R-sq.	0.09	0.11

Note: Values in parentheses are standard errors.
*Significant at 10% **Significant at 5% ***Significant at 1%.

people go in pairs under the non-monitoring condition, they have an opportunity to cheat and share the spoils. However, this does not happen. Variation in individual response across individual and group behavior is similar (Figures 6a and b). In both cases a small proportion of the sample over-reports while a much higher proportion of the sample under-reports. In the non-monitoring case and when people go alone, they under-report by an average of 9.7 percent while when they go in pairs, they under-report by an average of 8.9 percent (Figure 7). The two percentages are very similar and not significantly different. As we discussed in the previous section, we believe that this under-reporting is merely a random error unless pertaining to collection of cash incentives. In any case, the overall outcome of under-reporting even in groups

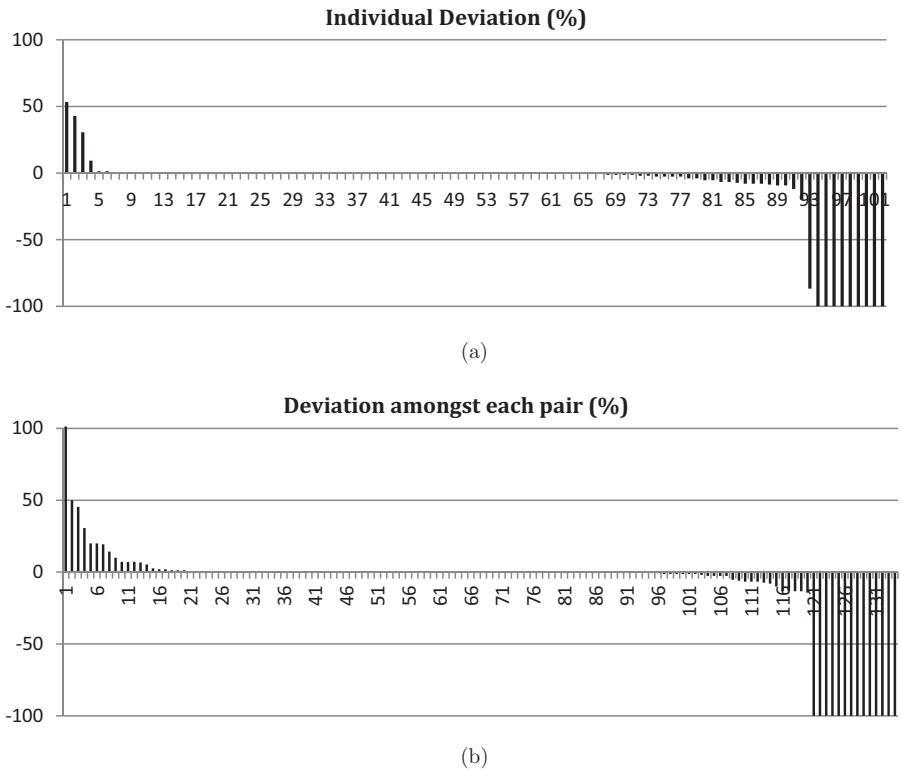

Figure 6: Variation in individual responses. (a) Deviation when people go alone to collect their compensation (b) Deviation when people go in pair to collect their compensation.

brings us to our second result:

People do not necessarily exhibit dishonest behavior when operating in groups versus when they are alone.

Result 3: Even when honest, cash, and in-kind incentives lead to different outcomes

Although, we do not find an increase in cheating with in-kind benefits, we do find a significant difference in behavior when people face cash versus in-kind incentives (Heyman and Ariely, 2004). Overall, the level

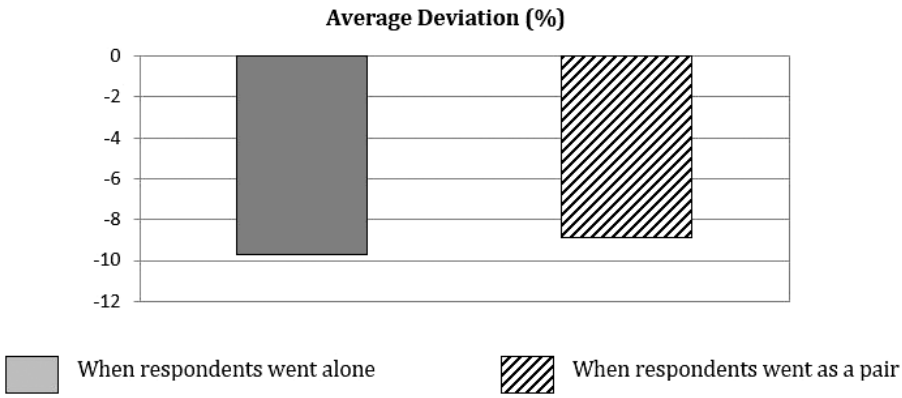

Figure 7: Average deviation in compensation collected (in percentage) across all respondents when they went alone or as a pair.

of under-reporting is much higher for cash than for in-kind incentives (Figure 9). When people are not monitored and when cash incentive is offered, a small sub-sample over-reports but a significantly larger proportion of the sample under-reports (Figure 8a). However, in case of seeds, the proportion of the sample that over-reports is roughly similar to the one that under-reports (Figure 8b). Overall, the average deviation is –3.4 percent in case of seeds. However, the magnitude of under-reporting is much higher and significantly different at –14.9 percent when people collect cash (Figure 10). This suggests that:

Even when people are honest, they are more cautious (stricter) in dealing with cash than with in-kind incentives.

Result 4: When honest, people are stricter on others

Dana *et al.* (2007) report that when behaving opportunistically, people are flexible about bending rules for themselves but not for others. Our experimental data supports this hypothesis but in a different way. When people are scrupulously honest, they do not bend rules for themselves, but they can be stricter on others than on themselves. In our experiment, when people collect compensation (whether cash or in-kind) for themselves, the proportion that over-reports is similar to

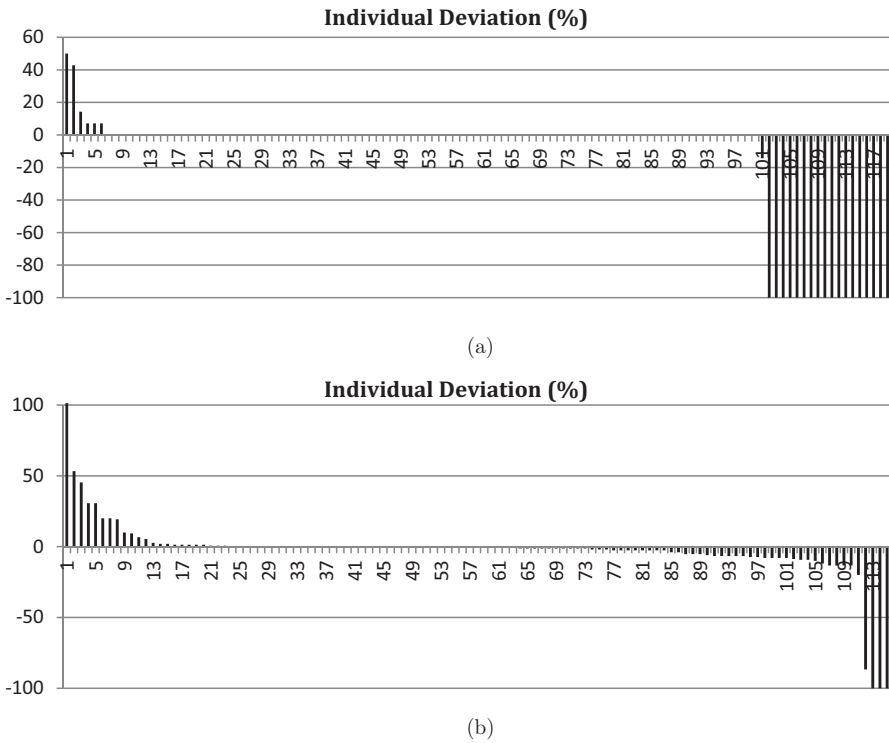

Figure 8: Variation in response to cash versus in-kind incentive. (a) Variation in individual responses when offered cash (b) Variation in individual responses when offered seeds.

the proportion that under-reports (Figure 10a), which is more a random error than a display of strategic behavior. However, there is a marked change when people collect compensation for their neighbors. Now the proportion of people under-reporting is much larger than the proportion that over-reports (Figure 10b). Further, out of the 118 times when subjects collect incentive for others while not being monitored, they under-report in 35 cases (29.7 percent of cases).

But how does this behavior vary depending on the kind of incentive offered? Figure 11 shows that there is no difference in average deviations when people collect cash or seeds for themselves, in fact the level of incentive collected is very close to the recommended level (average deviation of –1.19 percent for cash and –0.7 percent for seeds). However,

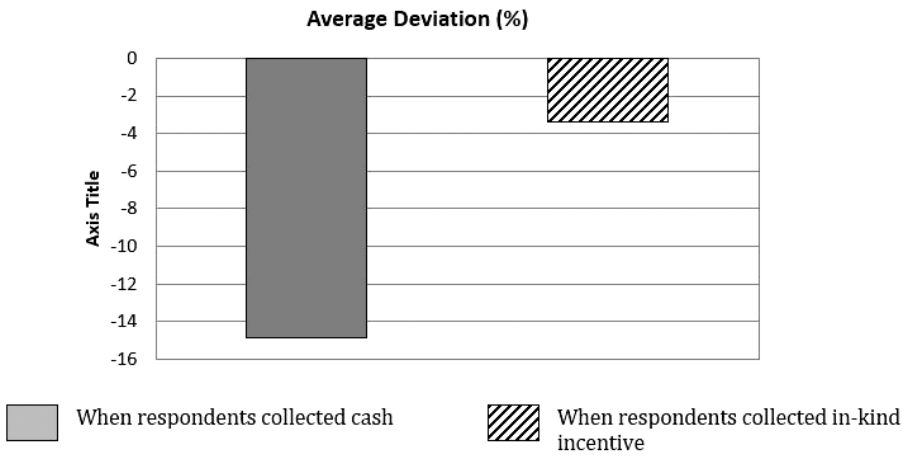

Figure 9: Average deviation in compensation collected (in percentage) across all respondents when they were offered cash or in-kind incentive.

when collecting incentive for their neighbors, they under-report by an average of –6.1 percent in case of seeds, and by a significantly larger –28.6 percent in case of cash (Figure 11).

We isolate this interaction between type of incentive and whether it is collected for self or for others by adding another variable to regression model (1). Now,

$$Y_i = a + b_1 X_{i1} + b_2 X_{i2} + b_3 X_{i3} + b_4 X_{i4} + b_5 X_{i5} + e_i \qquad (2)$$

where all the other variables are same as before and the new variable:

X_{i4} = Interaction between kind of incentive and whether or not it is collected for self or others (kind of incentive*collected for self or others)

= 1 when cash incentive is collected for others (cash*others), and

= 0 otherwise

Table 4 (column 2) presents the result of the OLS regression for model 2. The signs of various coefficients are the same as before. The interaction between kind of incentive offered and whether it is collected for others

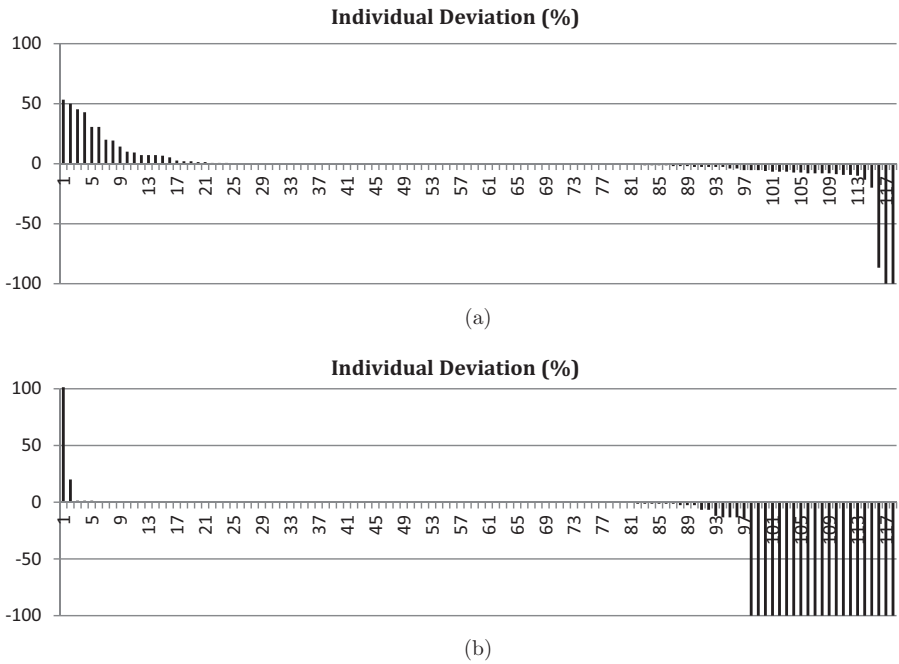

Figure 10: Variation in individual responses when people collect compensation for self versus for others. (a) Variation when respondents collect incentive for self (b) Variation in responses when respondents collect incentives for others.

is significant and negative (coefficient –17.8, *t*-stat 2.89). However, the difference in response due to cash and seeds is no longer significant (coefficient –1.2, *t*-stat 0.26). The entire difference that we see in response to cash and seeds (column 1, coefficient –10.1) is now captured by the interaction between incentive type and who it is collected for. This leads us to our final result:

When honest, people are stricter on others, especially when cash is involved.

4.3 Discussion

These results indicate predominantly honest behavior among our subjects, which is in line with findings of Pascual-Ezama *et al.* (2015), but strikingly different from those observed by Mazar *et al.* (2008), and

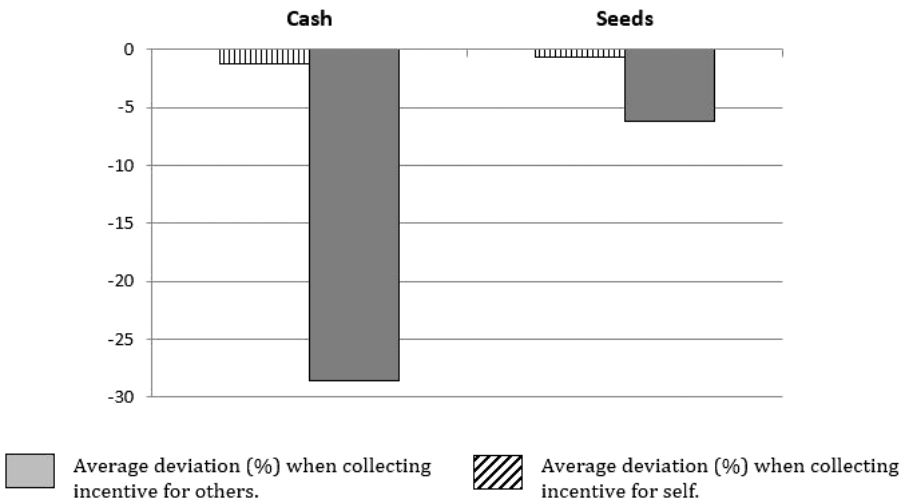

Figure 11: Average deviation in cash or seeds collected (in percentage) when respondents collected the incentive for themselves or for others.

those cited by Ariely (2012). As we discussed in Section 2, many of these studies refer to the concept of moral balancing or moral wiggle room that involves people feeling cognitive dissonance when their actions do not match with prevailing norms (Spiekermann and Weiss, 2013). To address this dissonance, people change how they perceive their actions, but not their actions themselves. Thus, they end up cheating even though they convince themselves that they are behaving honestly (Mazar and Ariely, 2006). However, many of the studies that refer to this moral balancing do not account for contextual factors that may affect how respondents behave (e.g. see Mazar *et al.*, 2008). This may be due to their experimental evidence being mostly collected through lab experiments. We are not questioning the generalizability of lab results, but merely pointing out that perhaps changing the context under which subjects operate could significantly change the experimental outcomes.

In our case, both the survey data and the field experiments point towards high level of honesty. Whenever over-reporting of incentives takes place, it is more due to random weighing error than strategic manipulation. In addition, the survey data shows that 67 percent of our respondents feel that people in their community will collect the

correct level of incentive because they are honest, and another 30 percent believe that cheating is unlikely to happen because people will follow instructions. When viewed along with the high level of trust and helping behavior as documented by the survey, there is strong reason to believe that cheating is not ubiquitous. Previous research has mainly focused on how moral wiggle room expands to justify cheating (Ploner and Regner, 2013), but not on the possibility of its narrowing under the influence of contextual factors such as prevalence of strict rule abiding in a society and high levels of trust and honesty.

Further, even though most survey respondents predicted that their community members will be honest, only a small proportion could foresee the possibility of them being more cautious when collecting incentive for their neighbors. However, during experiments, in the self versus others treatment, a significantly large number of people under-reported incentives for their neighbors, with many not leaving any cash at all. We do not think that this happened because these subjects did not wish their neighbors to be compensated. Rather, they seem to have become more cautious, which did not happen when they collected in-kind incentives for their neighbors. This finding provides evidence to the assertion that when people internalize social norms of trust and honesty, they do not suffer from cognitive dissonance when tempted to cheat. As their moral wiggle room becomes narrower, they continue to behave honestly even when they are not being observed by others. However, they may become stricter when monitoring others, especially when dealing with cash. The theoretical and practical implications of these results are discussed in the concluding section below.

5 Overall Discussion and Conclusion

The central question that we ask in this paper is do people cheat when they are unsupervised? Our interest in this question arises out of its close connection with behavioral aspects of self-monitoring in environmental management programs such as PES. To address this question, we carried out two field studies in rural Vietnam: a survey presenting hypothetical PES-like scenarios to 157 randomly selected respondents across five villages, and a set of behavioral experiments testing the same scenarios as treatments with 243 subjects from the

same area. The experiment (and the hypothetical scenarios in the survey) involved asking respondents to fill out a questionnaire on local agricultural and forestry practices, at the end of which they are asked to collect their compensation for their time. Treatments vary by whether the subjects were monitored or not when they collect their compensation, whether they received cash or in-kind incentives, whether they went individually or in pairs to collect their compensation, and how much did they collect for themselves versus for a resident of a neighboring village. In each case, the amount of compensation that was made available was more than what the respondents were asked to collect, thus providing them with an opportunity to cheat.

Previous work on human behavior in the absence of external monitoring has either taken the form of a deterministic model that predicts cheating whenever benefits supersede costs (Becker, 1968; Polinsky and Shavell, 1984), or a set of lab experiments that predict cheating by most people although of a lower magnitude (Ariely, 2012; Mazar *et al.*, 2008; Wiltermuth, 2011). In our two studies, we find very little evidence of cheating. This result corresponds closely to the results of research done by Pascual-Ezama *et al.* (2015). In our study, most subjects collect the correct level of compensation, which is also predicted by the survey responses. In case of in-kind incentive of maize seeds, there is evidence of a random weighing error that results in slight over-reporting as well as under-reporting with an average deviation that is not significantly different from zero. In the case of cash compensation, especially when subjects collect money for their neighbors, many respondents become overly cautious and do not leave any money at all. When these results are viewed along with the survey results, it becomes clear that when people are honest and law abiding, they may become even more careful when they are monitoring others.

We believe that our experimental results have both theoretical and practical significance. Our work shows that many existing models of human behavior are unable to predict the outcomes that we have observed. Results from our experiments in the field differ significantly from previous lab-based work that shows that participants' context does not play a role in explaining behavior (Mazar *et al.*, 2008; Pascual-Ezama *et al.*, 2015). In contrast, we find context as an important explanatory variable. Participants that come from an area with high levels of rule abiding behavior, and existence of strong norms of trust

and cooperation, exhibit high levels of honesty. Methodologically, our results emphasize the need to understand the study context and the importance of conducting behavioral experiments in the field rather than confining them to lab settings alone (Harrison and List, 2004).

Moreover, our work builds further on the theory of self-concept maintenance as proposed by Ariely (2012), Mazar *et al.* (2008), and Mazar and Ariely (2006). We find that instead of broadening, moral wiggle room may become narrower in societies with strong norms around honesty, trust, and law abiding behavior. In such cases, subjects face little distress in choosing honest behavior even when they are not being watched by others. The calculus of benefit–cost analysis, or of arguing with their inner self to cheat a little bit does not seem to arise. In addition, these honest and rule-abiding people do not bend rules for themselves but may end up being stricter monitors of others. As far as we know, this narrowing of moral wiggle room when base rules for moral behavior are extremely high, hasn't been explored by experimental literature and could become the focus of future studies.

Further, even though most survey respondents in our study correctly predicted that their community members will not cheat, they could not foresee the possibility of incorrect weighing of in-kind incentives. This happened because survey respondents assumed the everyday task of weighing seeds as simple, even though in practice, it became a challenge for a significant number of participants. We believe that this outcome has important implications for environmental programs. Local farmers may perceive conservation activities as simple and manageable, but in practice they may differ from the required technical standards. In this case, the importance of training community members and demonstrating how to take correct measurements in the field cannot be overstated (Holck, 2008; Skutsch, 2012). Most survey respondents, however, did not foresee the possibility of their neighbors being too cautious when collecting incentive for others. Only 1 percent of survey respondents felt that their neighbors would be stricter on others. In contrast, during field experiments, under-reporting for others was observed in almost 30 percent of the non-monitoring cases. The significant difference between survey responses and experimental behavior indicates that when people see themselves as honest and law abiding, they do not realize that they may be following very strict standards when monitoring others. This is another finding that needs to be explored more in future studies.

In terms of practical application, our study shows that local context and social norms around honesty and rule abidance are important in designing monitoring arrangements for environmental programs. This becomes especially relevant, given the growing number of PES-based community forestry projects that are being taken up in different parts of the world under the Paris Climate Agreement.

Before concluding, we would also like to point some important limitations that affect our analysis. Even though we would have liked to design treatments along typical PES activities (such as being compensated in cash or in-kind for planting trees), this was not possible due to time and resource constraints. Instead, we opted for PES-like scenarios such as being compensated for filling a questionnaire, which is often the first stage of many environmental projects in developing countries. However, some of our results may not conform to what may happen in actual PES projects particularly that involve repeated interactions which provide participants with an opportunity to learn over time. These limitations can be addressed by future studies that design more PES-based treatments and conduct a longitudinal analysis of people's behavior over time. Finally, our survey results may have been potentially affected by prevalence of another social norm of not saying anything bad about fellow villagers. This would be difficult to differentiate from the norm of people being honest as we document in our study. However, based on consistently correct collection of compensation by a large segment of our experimental subjects, we are confident in saying that our respondent group indeed followed strong norms of trust and honesty.

Acknowledgements

We thank participants at seminars organized by the University of Ottawa, the University of Toronto, and the World Agroforestry Centre for their comments and suggestions. Special thanks to Mamta Vardhan for helping to edit the main text. We also acknowledge the grant support from USAID Linkage Funds for carrying out this research, and the additional support that we received from the Social Sciences and Humanities Research Council of Canada (SSHRC) under its Banting Fellowship program. Helpful feedback from two anonymous reviewers is also gratefully acknowledged.

References

Ariely, D. (2012), *The (Honest) Truth About Dishonesty: How We Lie to Everyone — Especially Ourselves*, Harper Collins, Inc.

Bazerman, M. H., G. Loewenstein, and D. A. Moore (2002), "Why Good Accountants Do Bad Audits", *Harvard Business Review*, 80(11), 96–103.

Becker, G. S. (1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy*, 76(2), 169–217.

Bicchieri, C. (2008), "The Fragility of Fairness: An Experimental Investigation on the Conditional Status of Pro-social Norms", *Philosophical Issues*, 18(1), 229–48.

Bouma, J., E. Bulte, and D. van Soest (2008), "Trust and Cooperation: Social Capital and Community Resource Management", *Journal of Environmental Economics and Management*, 56(2), 155–66.

Brent, D. A., L. Friesen, L. Gangadharan, and A. Leibbrandt (2017), "Behavioral Insights from Field Experiments in Environmental Economics", *International Review of Environmental and Resource Economics*, 10(2), 95–143.

Bucki, M., D. Cuypers, P. Mayaux, F. Achard, C. Estreguil, and G. Grassi (2012), "Assessing REDD+ Performance of Countries with Low Monitoring Capacities: The Matrix Approach", *Environmental Research Letters*, 7(1), 1–13.

Campbell, E. Q. (1964), "The Internalization of Moral Norms", *Sociometry*, 391–412.

Cárdenas, J.-C. and E. Ostrom (2004), "What Do People Bring into the Game? Experiments in the Field About Cooperation in the Commons", *Agricultural Systems*, 82(3), 307–26.

Carpenter, J. P., A. G. Daniere, and L. M. Takahashi (2004), "Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums", *Journal of Economic Behavior & Organization*, 55(4), 533–51.

Cleaver, F. (2000), "Moral Ecological Rationality, Institutions and the Management of Common Property Resources", *Development and Change*, 31(2), 361–83.

Dana, J., R. A. Weber, and J. X. Kuang (2007), "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness", *Economic Theory*, 33(1), 67–80.

Danielsen, F., M. Skutsch, N. D. Burgess, P. M. Jensen, H. Andrianandrasana, B. Karky, and R. Lewis (2011), "At the Heart of REDD+: a Role for Local People in Monitoring Forests?", *Conservation Letters*, 4(2), 158–67.

Epley, N. and D. Dunning (2000), "Feeling Holier than Thou: Are Self-Serving Assessments Produced by Errors in Self-or Social Prediction?", *Journal of Personality and Social Psychology*, 79(6), 861–75.

Epley, N. and D. Dunning (2006), "The Mixed Blessings of Self-Knowledge in Behavioral Prediction: Enhanced Discrimination but Exacerbated Bias", *Personality and Social Psychology Bulletin*, 32(5), 641–55.

Fisher, R. J. (1993), "Social Desirability Bias and the Validity of Indirect Questioning", *Journal of Consumer Research*, 20(2), 303–15.

Fisman, R. and E. Miguel (2007), "Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets", *Journal of Political Economy*, 115(6), 1020–48.

Harrison, G. W. and J. A. List (2004), "Field Experiments", *Journal of Economic Literature*, 42(4), 1009–55.

Heyman, J. and D. Ariely (2004), "Effort for Payment a Tale of Two Markets", *Psychological Science*, 15(11), 787–93.

Hoang, M. H., T. H. Do, M. T. Pham, M. van Noordwijk, and P. A. Minang (2013), "Benefit Distribution Across Scales to Reduce Emissions from Deforestation and Forest Degradation (REDD+) in Vietnam", *Land Use Policy*, 31, 48–60.

Holck, M. H. (2008), "Participatory Forest Monitoring: An Assessment of the Accuracy of Simple Cost–Effective Methods", *Biodiversity and Conservation*, 17(8), 2023–36.

Jindal, R., J. Kerr, T. Dillaha, and M. Colby (2007), "USAID PES Sourcebook: Lessons and Best Practices for Propoor Payment for Ecosystem Services", OIRED, USA. (http://www.katoombagroup.org/documents/tools/PES.Sourcebook.PDF.pdf).

Kenney, L., J. Arvai, M. Vardhan, and D. Catacutan (2015), "Bringing Stakeholder Values into Climate Risk Management Programs: Decision Aiding for REDD in Vietnam", *Society and Natural Resources*, 28(3), 261–79.

Kerr, J., M. Vardhan, and R. Jindal (2012), "Prosocial Behavior and Incentives: Evidence from Field Experiments in Rural Mexico and Tanzania", *Ecological Economics*, 73, 220–7.

Lopez, M. C., J. J. Murphy, J. M. Spraggon, and J. K. Stranlund (2012), "Comparing the Effectiveness of Regulation and Pro-Social Emotions to Enhance Cooperation: Experimental Evidence from Fishing Communities in Colombia", *Economic Inquiry*, 50(1), 131–42.

Lusk, J. L. and F. B. Norwood (2009a), "An Inferred Valuation Method", *Land Economics*, 85(3), 500–14.

Lusk, J. L. and F. B. Norwood (2009b), "Bridging the Gap between Laboratory Experiments and Naturally Occurring Markets: An Inferred Valuation Method", *Journal of Environmental Economics and Management*, 58(2), 236–50.

Mazar, N., O. Amir, and D. Ariely (2008), "The Dishonesty of Honest People: A Theory of Self-concept Maintenance", *Journal of Marketing Research*, 45(6), 633–44.

Mazar, N. and D. Ariely (2006), "Dishonesty in Everyday Life and its Policy Implications", *Journal of Public Policy & Marketing*, 25(1), 117–26.

Moore, D. A., P. E. Tetlock, L. Tanlu, and M. H. Bazerman (2006), "Conflicts of Interest and the Case of Auditor Independence: Moral Seduction and Strategic Issue Cycling", *Academy of Management Review*, 31(1), 10–29.

Nguyen, Q., M. C. Villeval, and H. Xu (2012), "Trust and Trustworthiness Under the Prospect Theory: A Field Experiment in Vietnam", Available at SSRN 2144628 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2144628).

Ostrom, E. (2000), "Collective Action and the Evolution of Social Noms", *The Journal of Economic Perspectives*, 14(3), 137–58.

Parks, C. D. and A. D. Vu (1994), "Social Dilemma Behavior of Individuals from Highly Individualist and Collectivist Cultures", *Journal of Conflict Resolution*, 38(4), 708–18.

Pascual-Ezama, D., T. R. Fosgaard, J. C. Cardenas, P. Kujal, R. Veszteg, B. G.-G. de Liaño, B. Gunia, D. Weichselbaumer, K. Hilken, A. Antinyan, J. Delnoij, M. D. Tira, Y. Pratomo, T. Jaber-López, and P. Brañas-Garza (2015), "Context-Dependent Cheating: Experimental Evidence from 16 Countries", *Journal of Economic Behavior & Organization*, 116, 379–86.

Ploner, M. and T. Regner (2013), "Self-image and Moral Balancing: An Experimental Analysis", *Journal of Economic Behavior & Organization*, 93, 374–83.

Polinsky, A. M. and S. Shavell (1984), "The Optimal Use of Fines and Imprisonment", *Journal of Public Economics*, 24(1), 89–99.

Price, M. K. (2014), "Using Field Experiments to Address Environmental Externalities and Resource Scarcity: Major Lessons Learned and New Directions for Future Research", *Oxford Review of Economic Policy*, 30(4), 621–38.

Qin, X., J. Shen, and X. Meng (2011), "Group-Based Trust, Trustworthiness and Voluntary Cooperation: Evidence From Experimental and Survey Data in China", *The Journal of Socio-Economics*, 40(4), 356–63.

Scarpa, R., E. S. K. Ruto, P. Kristjanson, M. Radeny, A. G. Drucker, and J. E. O. Rege (2003), "Valuing Indigenous Cattle Breeds in Kenya: An Empirical Comparison of Stated and Revealed Preference Value Estimates", *Ecological Economics*, 45(3), 409–26.

Skutsch, M. (2012), *Community Forest Monitoring for the Carbon Market: Opportunities under REDD*, Routledge.

Spiekermann, K. and A. Weiss (2013), "Objective and Subjective Compliance: How Moral Wiggle Room Opens", Department of Government, The London School of Economics and Political Science. (http://personal.lse.ac.uk/SPIEKERK/papers/Spiekermann_Weiss_2013-10-17_web.pdf).

Steblay, N. M. (1987), "Helping Behavior in Rural and Urban Environments: A Meta-Analysis", *Psychological Bulletin*, 102(3), 346–56.

Wilson, R. S., J. L. Arvai, and H. R. Arkes (2008), "My Loss is Your Loss . . . Sometimes: Loss Aversion and the Effect of Motivational Biases", *Risk Analysis*, 28(4), 929–38.

Wiltermuth, S. S. (2011), "Cheating more When the Spoils are Split", *Organizational Behavior and Human Decision Processes*, 115(2), 157–68.

Wong, G., A. Angelsen, M. Brockhaus, R. Carmenta, A. Duchelle, S. Leonard, C. Luttrell, C. Martius, and S. Wunder (2016), "Results-based Payments for REDD+: Lessons on Finance, Performance, and Non-Carbon Benefits", CIFOR Infobrief 138, Center for International Forestry Research (CIFOR), Bogor, Indonesia. (http://www.cifor.org/publications/pdf_files/infobrief/6108-infobrief.pdf).