Incentives and Social Preferences:

Experimental Evidence from a Traditional Labor Contract[*]

Jun Goto
Hitotsubashi University

Takeshi Aida
Graduate National Research Institute for Policy Studies

Keitaro Aoyagi
University of Tokyo

Yasuyuki Sawada[1]
University of Tokyo

November 2014

Abstract

This paper investigates the interplay between economic incentives and social norms in formulating rice planting contracts of the Philippines. In our study area, despite the potential of infestation of opportunistic behaviors by workers, a fixed wage (FW) contract has been dominant for rice planting since the 1960s. To account for this puzzle of a seemingly-inefficient contractual arrangement, we adopt a hybrid experimental method of framed field experiments by randomly assigning three distinct labor contracts, i.e., FW, individual piece rate (IPR), and group piece rate (GPR) contracts and artefactual filed experiments to elicit social preference parameters. Through the analyses of individual workers' performance data from framed field experiments and data on social preferences elicited by artefactual field experiments, Three main empirical findings emerge. First, our basic results show the positive incentive effects in IPR and, equivalently, moral hazard problems in FW, which are consistent with standard theoretical implications. Second, non-monetary incentives seem to play a significant role under FW: while social preferences such as altruism and guilt aversion play an important role in stimulating incentives under FW, introducing monetary incentives crowds out such intrinsic motivations; and other non-monetary factors such as positive peer effects significantly enhance incentives under FW contract. Finally, as alternative hypotheses, our empirical results are not necessarily consistent with the hypothesis of the interlinked contract of labor and credit transactions in mitigating moral hazard problems, the optimality of FW contract under large effort measurement errors, and the intertemporal incentives arising from performance-based contract renewal probabilities. Hence, considering the interplay of intrinsic motivations and monetary incentives, we may conclude that seemingly perverse traditional contractual arrangements might not be too inefficient.

Keywords: Incentives, social preferences, peer effect, labor contract, field experiment

JEL: D03, C93, D22, C91

---

[*] The authors would like to thank Jonna P. Estudillo, Thelma B. Estera, Hitoshi Matsushima, Tetsuji Okazaki, Keijiro Otsuka, Hideo Owan, and Tomomi Tanaka for their encouragements to initiate this project and Oriana Bandiera, Alan de Brauw, Alain de Janvry, Ronald Findley, Tomoki Fujii, Michael Carter, Julen Esteban-Pretel, Marcel Fafchamps, Drew Griffen, Hal Hill, Hide Ichimura, Shingo Ishiguro, Michihiro Kandori, Takashi Kurosaki, Chiaki Moriguchi, Akihiko Matsui, Alistair Munro, Valerie Mueller, Takeshi Murooka, Masayuki Nakabayashi, Ryo Nakajima, Peter Orazem, Keijiro Otsuka, Betty Sadoulet, Tatsuyoshi Saijo, Paul T. Schultz, Andrew Sim, Katsuya Takii, Eric Thorbecke, Angelino Viceisza, Eric Weese, David Yanagizawa-Drott, Jane Zhang, and participants of MIEDC 2012 at University of Minnesota, NEUDC 2011 at Yale University, International Food Policy Research Institute, Social Science Division of International Rice Research Institute, 2012 and 2013 Asia-Pacific ESA Conference at Xiamen University and in Tokyo, respectively, ICAE 2012, and Australian National University, GRIPS, Kyoto University, Singapore Management University, UC Davis, University of London, University of Tokyo, Waseda University and Yokohama National University for useful comments on earlier drafts. The usual disclaimers apply.

[1] Corresponding Author: Faculty of Economics, University of Tokyo; Email: sawada@e.u-tokyo.ac.jp; Tel: +81-3-5841-5572; Address: 7-3-1 Bunkyo-ku, Tokyo 113-8657 Japan

1. Introduction

In the last forty years, the fixed wage (FW) contract for rice planters has been the dominant contract in the Central Luzon province of the Philippines. Since the supervisor of rice planting called *kabisilya* can observe individual work-effort outcome quite accurately to some extent, such an FW contract with low intensity of incentives is likely to be sub-optimal, generating a moral hazard problem. In this paper, we aim to solve this puzzle of the real-world perverse contract by a hybrid experimental method of framed field experiments by randomly assigning different labor contracts and artefactual filed experiments to elicit social preference parameters.[2] By doing so, we believe that we will be able to better understand economic rational of existing seemingly-inefficient agricultural labor contracts which are crucial in setting rural development and poverty alleviation strategies (Hayami and Otsuka, 1993; Roumasset and Lee, 2007).

Largely speaking, there are two theoretical ways to argue the optimality of FW contracts. First, in line of the arguments by Bowles (2008), intrinsic motivations may prevent workers from taking opportunistic behavior even under a FW contract.[3] The importance of the interaction between intrinsic motivations and monetary incentives in the workplace has been incorporated in economic theory (Kreps, 1997; Benabou and Tirole, 2003; Itoh, 2004; List and Rasul, 2011; Bowles and Polania-Reyes, 2012).[4] This attempt extends standard agency models to take into account other-regarding preferences in mitigating moral hazard problems (Kandel and Lazear, 1992; Rotemberg, 1994; List and Rasul, 2011).

Second, with standard theories, we may be able to explain the optimality of FW contracts. First, since the optimal intensity of incentive condition in a standard agency model tells us that the slope of performance-based wage compensation scheme is a negative function of observability or precision

[2] While in each day of field experiments, we work with basically the same set of workers, assigning randomly different contract in the morning and the afternoon sessions, participating rice planters differ across days. This setting provide us with within-subject variations as well as between subject variations in our data.

[3] In fact, such intrinsic motivations may arise from the Filipino community norms such as *pakikisama*, literally meaning "going along well with others," (Hayami and Kikuchi, 2000).

[4] Itoh (2004) calls this new research field "behavioral contract theory."

of worker performance (Milgrom and Roberts, 1992), with sufficiently large measurement errors of worker performance, it is optimal for the principal to set a FW contract. Secondly, the optimality of FW contracts can be shown in a multitask principal-agent model in which performance of one of the indispensable tasks is unmeasurable or unverifiable (Proposition 1 of Holmstrom and Milgrom, 1991). While it may be said theoretically that rice planting is a multitask work of planting speed and adequacy, in reality, rice planting tasks are rather simple and are difficult to manipulate (Umekage at al., 1965).[5] Third, the intertemporal incentive for contract renewal can also be consistent with the optimality of a FW contract. In a situation where a supervisor has a right to decide contract renewal of workers, contract renewal probability can be a positive function of work performance which can be observed to some extent. Then, while instantaneous wage is fixed, expected wage becomes piece rate effectively, mitigating moral hazard. Finally, the FW contract can be optimal if it is interlinked with other contract such as a credit contract. This is simply because by changing the terms and amount of the loans to the worker, lenders can induce more efforts, solving moral hazard problems in labor contracts (Braverman and Stiglitz, 1982).

As to the first way to explain the optimality of FW contracts, some recent empirical studies using non-experimental personnel data suggest that "social incentives," which are an important part of other-regarding preferences, can increase the worker's productivity and solve agency problems in a FW contract (Mas and Moretti, 2009).[6] Yet, the results reported in these previous empirical studies may involve estimation biases because observed modes of contracts might well be endogenous to firm's performance and deliberate choices (Prendergast, 1999). Hence, there seems to be a room for

[5] Umekage et al. (1966) carried out experiments to identify the effect of different transplanting methods on the growth habit of the rice plant by comparing two methods, i.e., planting the rice seedlings in the horizontal direction and in the vertical direction which is the normal method. They found that the difference of yield between the plots was not significant at the 5% level. Moreover, according to the supervisor of the rice planting workers called *kabisilyas*, not only speed but also inadequacy of planted rice can be easily observed and verified by checking misalignments and lodging of seedlings. Hence, we believe that multi-task problems are not necessarily serious in our experiments. Also, in our instruction of experiments, we emphasize that we will not make experiment payments for lodged seedlings so that we can mitigate the multi-task problems. *Kabisilya* told us that there are no systematic negative harvest effects arising from possible lodging of seedlings.

[6] Bandiera et al. (2010) also find that social incentives affect incentives of workers significantly.

improving identification of causal impacts of a particular incentive scheme on individual performance. Furthermore, individuals may make decisions to choose their peers, making it difficult to determine the causality from a contract to worker behavior. In sum, we may safely say that what sort of social motivation shapes people's behavior and how these motives interact with monetary incentives are still under-investigated in the literature.

The objective of this paper is to bridge this gap in the existing empirical studies in two ways. First, we investigate the interplay between economic incentives and intrinsic motivations. To be more precise, we examine whether monetary incentives crowd out intrinsic motivation by a hybrid experimental method of artefacutual and framed field experiments (Harrison and List, 2004; Levitt and List, 2009). Second, we test alternative hypotheses to the crowding out hypothesis, i.e., the effort measurement error, intertempotal incentives, and interlinked contract hypotheses. To these purposes, we combine an individual performance data from rice planting framed field experiments in the Central Luzon Province of the Philippines with the data on social preferences constructed from the results of artefactual field experiments, which enable us to specify the underlying mechanisms of mitigating agency problems. In these field experiments, we randomly assign three distinct labor contracts, i.e., FW contract as well as individual piece rate (IPR) and group piece rate (GPR) contracts.

To preview our empirical results, three main findings emerge. First, our basic results show the positive incentive effects in IPR and moral hazard problems in FW, which are consistent with standard theoretical implications as well as empirical findings by Lazear (2000). Second, non-monetary incentives seem to play a significant role under FW. While social preferences such as altruism and guilt aversion play an important role in stimulating incentives under FW, introducing monetary incentives crowds out such intrinsic motivations., a rare real world finding which confirms evidence from laboratory experiments reviewed by Bowles (2008). Also, other non-monetary factors such as positive peer effects significantly enhance incentives under FW contract, which is consistent with findings by Mas and Moretti (2009). Finally, as alternative hypotheses, our empirical results are not necessarily supportive to the hypothesis of the interlinked contract of labor and credit transactions in

mitigating moral hazard problems, the optimality of FW contract under large effort measurement errors, and intertemporal incentives arising from performance based contract renewal probabilities. In sum, while social preferences appeared as altruism and/or inequality aversion seem to mitigate agency problems in the workplace, these effects might have been crowded out by the monetary incentives. Hence, considering the interplay of intrinsic motivations and monetary incentives, we may conclude that this seemingly perverse traditional contractual arrangements might not be too inefficient.

The rest of this paper is organized as follows. Section 2 describes the context of our study and our framed and artefactual field experiments. In Section 3, we show our data with summary statistics which is followed by empirical results presented in Section 4. Section 5 summarizes concluding remarks.

2. Experimental Design

2.1 Rice Planting in the Study Villages

In the Central Luzon province of the Philippines, labor to transplant seedlings manually is traditionally supplied by agricultural laborers and members of tenant families from local villages at the beginning of each regular crop season. Since farmers do not want to miss the appropriate timing of rice planting when irrigation water becomes available, during peak planting season, demand for hired rice planting labor is concentrated around the same time in each village. To allocate planting labor effectively, loosely-tied labor groups for rice planting are formed by supervisors called *kabisilya*, literally meaning "boss." Figure 1 illustrates the structure of rice planting organization and payment scheme. First, a farmer outsource rice planting to a *kabisilya* by paying 2,200PHP/ha. The total amount of payment from a farmer to *kabisilya* only depends on total contracted size of paddy field. After *kabisilya* subtracts her fixed margin from the total revenue,[7] i.e., 2,200 PHP multiplied with total planting area size, the rest amount is equally divided among planting workers. While, theoretically

[7] The margin is 100-150 PHP.

speaking, number of planting workers is endogenously determined, the prevailing norm in the area was to recruit appropriate number of workers so that each daily payment becomes 100-120PHP per worker. In this way, workers receive fixed wage effectively regardless of their performance.[8] During the peak planting season, workers are asked to work on paddy fields almost every day and typically plant in one to two hector per day together with their peer workers. Workers are not permitted to plant in other landowner's plots before finishing their duties in a given day. In a paddy field, workers voluntarily form sub-groups that consist of five to seven members and take responsibility for planting in assigned plots. Composition of these sub-groups usually change daily. Accordingly, the choice variables left in hand for planting workers are selection of co-workers and the level of efforts put in planting.

2.2 A Hybrid Experiment

In this study, we adopt a hybrid method of framed field experiments and artefactual field experiments (Harrison and List, 2004; Levitt and List, 2009). The former experiment is to change rice planting labor contract; and the latter is to make laborers participate in a series of carefully-designed laboratory economic experiments.

2.2.1 Framed Field Experiments

In our framed field experiments, all workers are asked to take part in randomly assigned three contracts, i.e., FW, IPR, and GPR contracts. While group formation for actual rice planting has been based on self-selection mechanisms, we also introduce random group assignments. Hence, we have a total of six arms, i.e., combination of three assigned contracts and two ways of group formation, in our experiments. During the first half of the planting season, each planting group is formed randomly by

[8] Strictly speaking, the reward level for each worker relies on the number of planters. Yet, the ratio of the number of recruited workers in a day to total contracted area is hardly varied. Hence, the prevailing contract is pure fixed wage contract effectively.

us; and during the second half, workers form their own planting groups. At the beginning of the morning and afternoon planting of each day, we select a contractual arrangement randomly out of three contracts.[9] In each day, we work with the same set of workers, assigning randomly different contract in the morning and the afternoon sessions. This setting provide us with within-subject variations in our data. Yet, participating rice planters differ in each day. Hence, strictly speaking, we also utilize between subject variations.

Three Wage Contracts

Let l_{ij} to be the output level chosen by a worker *i* in a group *j*, that is measured by the average length of planting line in each ten minutes over thirty minutes.[10] Suppose that observed outcome is decomposed into the true effort level of worker *i*, e_{ij}, and the well-behaved measurement error, ε_{ij}, i.e., $l_{ij} = e_{ij} + \varepsilon_{ij}$. We assume that only l_{ij} can be observable and used for a performance-based contract. The prevailing rice planting contract is a FW contract under which the amount of fixed wage, w_{ij}^f, for a worker *i* belonging to sub-group *j* is simply: $w_{ij}^f = F$. Under IPR contract, a wage rate, w_{ij}^p, is composed of two parts –a fixed component and an incentive component which is proportional to individual productivity: $w_{ij}^p = F + \alpha(l_{ij})$, where $0 \leq \alpha(l_{ij}) \leq F$. Hence, the amount of rewards is set at the level that the most productive worker can get paid twice of fixed rate. Under GPR contract, a wage rate, w_{ij}^g, is also composed of two parts – a fixed compensation part and an incentive component which is proportional to average productivity of the group rather than each individual: $w_{ij}^g = F + \alpha\left(\sum_{i\in j} l_{ij}/n_j\right)$, where $0 \leq \alpha\left(\sum_{i\in j} l_{ij}/n_j\right) \leq F$ and n_j is the number of workers in a group *j*. Note that we set that the fixed wage, *F*, equals to fifty pesos and the intensity of incentive parameter, α, equals to two pesos.

It is straightforward to show that marginal monetary benefits of efforts in three contracts are:

[9] To avoid confounding the effects of each treatment with some other factors, the compensation schemes of kabisilya remains unchanged.

[10] While we measure the length for thirty minutes, we construct three observations of ten minutes data for each worker.

$\frac{\partial w_{ij}^{f}}{\partial e_{ij}} = 0$, $\frac{\partial w_{ij}^{p}}{\partial e_{ij}} = \alpha$, and $\frac{\partial w_{ij}^{g}}{\partial e_{ij}} = \frac{\alpha}{n_j}$, respectively. Hence, while the marginal monetary benefit is highest in IPR which is followed by GPR, the marginal monetary benefit is zero under FW. Hence, in terms of monetary incentive effects (Lazear, 2000), an observed gap in worker performance between IPR and FW reflects the moral hazard problem; and the difference of individual worker outcomes between IPR and GPR exhibits the free-riding problem.

Randomized Team Formation

The team formation procedure in our experiment is structured in two alternative ways: randomly-formed groups and self-selected groups. In both groupings, teams are newly formed for each rice planting session. The basic size of teams is a group of five workers but some adjustments are made based on the total workers who show up to the field.[11] In one setting, all planting workers recruited by a *kabisilya* at a given paddy field are randomly assigned into teams.[12] A plot is randomly assigned for each team by the experimenters. We can refer to this arrangement as "randomly-assigned groups." In the other setting, workers are asked to organize their planting teams by themselves in each session. We call this arrangement as the "self-selected groups."[13] Two potential factors drive team composition in the latter setting. First, all workers may prefer to match with high ability colleagues to complete planting in the assigned rice fields as quickly as possible. In this case, given other things equal, we would expect assortative matching patterns. Alternatively, workers may incline to form teams with socially connected members because they derive direct utility and/or they can achieve cooperation in a self-enforcing manner (Bandiera et al., 2005, 2010).

[11] The upper and lower bound of number of teams and their members are set due to the limited availability of plot for experiments and the securing accuracy in measuring productivity by enumerators.

[12] The number of workers recruited by *kabisilya* depends on the location of paddy field and total contracted area.

[13] Theoretically, the assignment of workers to formed teams represents a coalition proof Nash equilibrium (Bandiera et al, 2009, Berheim et al., 1987).

Procedures of Experiments

In each experiment, a contract type is randomly chosen and assigned by experimenters at the beginning of each experiment. This experiment is carried out twice in each day, i.e., in the morning and the afternoon. In order to avoid anticipation effects, a notice on the selected contract has been made to workers and *kabisilya* right before each rice planting session. Enumerators are also randomly assigned to each sub-group to measure workers' individual productivities and are required to record productivities by tracing each worker's planting line by handwriting on carefully-designed reporting sheets for the initial thirty minutes in each planting session.[14] Rewards for individual workers in each experiment are immediately paid at the paddy field before starting a new round of an experiment. The payments for each worker are kept strictly confidential so that other factors such as self-image construction are carefully controlled.

2.3 Artefactual field Experiments

To examine the optimality of FW contracts, we follow the arguments by Bowles (2008) to hypothesize that intrinsic motivations may prevent workers from taking opportunistic behavior even under a FW contract. Such intrinsic motivations may arise from the Filipino community norms such as *pakikisama*, literally meaning "going along well with others," (Hayami and Kikuchi, 2000). To elicit each worker's such social and individual preference parameters, we conducted four standard laboratory experiments with the same rice planting workers or "artefactual field experiments": dictator game, ultimatum game, public goods game with monitoring and disapproval message, and risk game. At the end of experiments, each subject received real monetary payoff based on a randomly-selected

[14] During the pilot experiments, this method has been validated by comparing it with alternative methods. First, we use video camera to recode each worker's all activities in a given plot; and second, we also ask workers to put high and low quality passometers on their waist during rice planting. Since video-taping and decoding are too costly and we cannot achieve consistent measurements using different passometers, we decided not to use them.

outcome of experiments.[15] We employ the strategy method for responder's behavior in ultimatum game to elicit punishment for unfair offers and fairness such as inequity aversion (Camerer and Fehr 2004; Levitt and List 2007).

Dictator Game

We conduct a standard dictator game in anonymous setting to elicit altruism toward a co-worker under the same *kabisilya* (Camerer and Fehr, 2004; Levittt and List, 2007). Each player was given an envelope with 10 coins of 10 PHP.[16] Then they decide how many coins to put his/her partner's envelop. However, none of the players were informed about identify of their partner.

Ultimatum Game to Elicit Guilt Aversion and Nonenviousness

The first move in ultimatum game is the same as the dictator game but the second move is added in this game (Roth, Prasnikar, Okuno-Fujiwara, and Zamir, 1991): Each sender or "proposer" is informed beforehand about the rule that their keeping and sending amounts materialize only if the "responder" accepts their offer; if the responder rejects the offered amount in the first move, both proposer and responder receive nothing. After each player decided how many coins to allocate to their partner, they were also asked which offer they would accept as a responder using the strategy method. The observed outcomes in ultimatum game can be explained using standard equilibrium concepts applied to the inequality aversion utility of Fehr and Schmidt (1999), in which players care about their own payoffs and the difference between their payoffs and those of others. Specifically, we quantify two aspects of inequality aversion. First, since the proposers' offer amounts in ultimatum game depend on the guilt weight of Camerer (2003, pp., 102-103) in the inequality aversion preference, we

[15] Detailed descriptions of artefactual field experiments and actual implementation procedures are available from the corresponding author upon request.

[16] 100 PHP is equivalent to the daily wage rate for manual labor. Each subject also receive show up fee of 150PHP and a free lunch. Payments for artefactual field experiments are made for a randomly chosen experiment out of all the experiments conducted.

interpret the sending amounts in ultimatum game as an observed level of "guilt aversion." Since altruism can also affect the sending amount, we analyze the data after subtracting the sending amount in dictator game from that in ultimatum game. Second, based on the inequality aversion model of Fehr and Schmidt (1999), the minimum acceptance level of a responder elicited in ultimatum game is a "negative" function of inequality aversion or the envy weight parameter of Camerer (2003). Hence, we take the gap between the maximum value, i.e., 10 coins of 10PHP, and the minimum acceptance level as a responder in ultimatum game as a measure of "nonenviousness."

Public Goods Game with Monitoring and/or Disapproval Messages

We conducted a public goods game with monitoring and/or disapproval messages based on Carpenter and Williams (2014). In this setting, workers formed anonymous groups of four where they stayed for the entire experiment. At the beginning of the first round, each player was given an envelope with 10 coins of 10 PHP. Then they were asked how many coins to contribute in the group project, keeping the rest for themselves in a non-cooperative setting. The total amount of group contribution was doubled and redistributed equally to all four members.

After this contribution stage, we showed each worker his/her gross income and total contribution. Then they were given a chance to monitor the contribution amounts of each group member by paying one PHP. If a worker monitored other players' contribution, then he/she could send messages of disapproval (unhappy faces) to other individuals in the group for an additional 1 PHP per message. We continued these contribution and monitoring stages for three rounds. However, they were not allowed to send any messages in the third round for which there is no room for improving other player's contribution level.

Since a pure-strategy Nash equilibrium of the public goods game is zero contribution for all subjects, the existing literature such as Camerer and Fehr (2004) and Levitt and List (2007) interpret observed contributions reflect a type of social preference due to *reciprocal expected cooperation*. Each worker's *propensity to monitoring* and *propensity to sanction* are captured by the raw monitoring

frequency and the number of times sending disapproval messages, respectively (Carpenter and Seki, 2005; and Carpenter and Williams, 2014).

Risk Game

In addition to these experiments eliciting social preference parameter, we also conducted a risky investment game of Schechter (2007) to measure each workers risk preference. Each subject was given an envelope with 10 coins of 10 PHP and decided how many coins to invest in a risky investment game and how many coins to take home without investments. Then he/she tossed a coin; if it was head, their invested amount was doubled, but if it was tail, the subject lost the entire invested amount. The invested amount in this game thus represents each worker's individual risk tolerance level.

3. The Data

Our two study villages, G village and M village, are located in Nueva Ecija Province in the Central Luzon region, Philippines. These villages were studied intensively by projects under Social Science Division of International Rice Research Institute (Otsuka, 1991; Otsuka, Chuma, and Hayami, 1993; David, Cordova, and Otsuka, 1994; Estudillo, Sawada, and Hossain, 2005; and Estudillo, Sawada, and Otsuka, 2008, 2009). We conduct framed and artefactual field experiments for ten and two days, respectively, during the dry planting season in 2011. Our sample is composed of 120 workers in which 58 and 62 workers are from G village and M village, respectively. Without prior notice to workers and enumerators, in each morning or afternoon session of each day, we randomly introduced one of the three distinct monetary incentives, i.e., FW, IPR, and GPR contracts, under random- and self-selection arrangements. We obtain a total 1,884 observations from 28 field experiment session, i.e., 15 sessions for G village and 13 sessions for M village over 10 days (Figure 2).

Additionally, we conducted three different surveys: First, a short questionnaire survey for workers after each field experiment session to capture about session-specific questions such as

subjective health conditions during the session; second, a household survey to collect information on individual and household characteristics of each worker, and, finally, a survey for *kabisilya* after the very last field experiment to evaluate subjective performance of each worker.

Towards the end of field experiments, we spent two days, one day for each village, to conduct artefactual field experiments: we invited the rice planting workers who participated in the rice planting experiments to play a series of laboratory experiments (Figure 2). In each day, the experiment took four to five hours. Total number of participants was 108 in which 53 and 55 were from G and M villages, respectively.

3.2 Descriptive Statistics of Data from Field Experiments

Table 1 shows descriptive statistics of individual rice planting performance in field experiments, the conditions of field experiments, and household characteristics collected by surveys. The panel A shows pro-incentive effects in the IPR contract: On average, workers in IPR could achieve the highest productivity (29.97 m/10 min) followed by GPR (29.75 m/10 min) and FW (26.08 m/10 min). While the average productivity difference between FW and GPR is statistically significant at 1% level, the mean difference between GPR and IPR is not significant. Figure 3 compares cumulative distribution functions (CDFs) of individual rice planting productivity in three contracts. We can easily verify that the CDF of the IPR contract dominates that of the FW contract. The two-sample Kolmogorov-Smirnov tests of the equality of distributions reject the equality between these two CDFs at the 1% level of statistical significance. In contrast, the equality hypothesis of CDFs between IPR and GPR contracts can be rejected only at the 10% significance level. This is not necessarily surprising because people often react to the mere presence of incentives rather than their extent (Bowles and Polania-Reyes, 2012).

To grasp the overall pattern on peer effects within each rice planting team, Figure 4 shows the correlation between individual productivity and other member's average productivity in the same team. We find that individual productivity is positively correlated with other's productivity, suggesting that

there is a strong positive peer effect.

The panel B of Table 1 presents the subjective relathionships among workers within each rice planting team. The table shows that the average years of aquintance with the group members is 10.64 years. Moreover, 39.38% of the group members go to the same church where regular services are provided on Sundays.

According to the panel C in Table 1 which summarizes individual characteristics, the average age of the workers is 35.6 years old. The subjects tend to be landless agricultural worker and almost all workers receive advanced payments or, equivalently, credit from *kabisilya* in every planting season. The average amount of outstanding advanced payment is 625 PHP. As to the pattern of self-reported relathionship with *kabisilya*, more than 38% of surveyed workers are family members or relatives of *kabisilya*. This is not necesarily unusual in Central Luzon where most workers come from the same *barangay*, i.e., village, of their *kabisilya* who usually recruit workers through her social network in the village.[18]

3.2 Descriptive Statistics of Data from Artrefactual Field Experiments

Panel A of Table 2 shows the summary statistics of the public goods game results. The mean contribution over rounds is 33.3 PHP out of 100 PHP, indicating that the workers contribute one-third of their initial endowment. Figure 5 is a histogram of each worker's mean contribution level to public goods over all rounds: most workers contribute less than 40% of their initial endowment. According to Table 2, there is no clear trend in contribution level over time, which is similar to the results of Carpenter and Williams (2014), but is different from the standard voluntary contribution games without monitoring and disapproval message (Ledyard 1995). As we can see from Table 2, the monitoring

[18] The anthropological literature describes barangay or more small unit of area like cluster as traditional community groups composed mainly of kin or extended family (Kerkvliet, 1990).

intensity decreases slightly as the round proceeds, a finding which is consistent with Carpenter and Williams (2014). The number of disapproval messages sent by each worker increases from the first round to the second round.[19]

Panel B of Table 2 shows the summary statistics of the dictator game and ultimatum game results. The average sending amounts in both games are consistent with the results reported in the previous studies such as Camerer and Fehr (2004) and Levitt and List (2007). The average sending amount in the ultimatum game is larger than that in the dictator game which is also consistent with the previous studies (Forsythe et al. 1994). The difference in the sending amount between the ultimatum games and dictator and can be interpreted as the degree of "pure" guilt aversion. Figure 6 and 7 show the histograms of the dictator and ultimatum games, respectively. While, in the dictator game, most of the workers send less than 50 PHP, there is a clear peak at twenty and fifty PHP in the ultimatum game. Figure 8 shows the histogram of the responder's minimum acceptance level in the ultimatum game. While the highest peak can be found at zero, the second highest peak is at fifty, impling the existence of the 50-50 norm (Andreoni and Bernheim, 2009).

Panel C of table 2 shows the result of the risk game. The mean invested amount is 37.7 and all workers invest at least ten PHP. Since the expected net return from this experiment is zero, the investment amount represents each worker's preference toward risks. Figure 9 shows the histogram of the risk game in which we can see that most investment amounts are less than or equal to fifty PHP.

4. Empirical Analysis

4.1 Econometric Framework

We follow Lazear (2000), Shearer (2004), and Mas and Moretti (2009) to postulate the following

[19] Note that it was not allowed to send any messages in the final round.

econometric model of rice planter's observed work performance:

$$l_{ijt} = \alpha_t + \lambda^{FW} FW_t + \lambda^{GPR} GPR_t + S_{ijt}\delta + \mu P_{ijt} + \pi R_t + X_{ijt}\gamma + \varepsilon_{ijt} \quad (1)$$

where l_{ijt} is the productivity of worker *i* in team *j*, measured in the average planting length in meters per ten minutes, in *t*-th round; α_t is round (time) effect; FW_t and GPR_t are dummy variables for FW and GPR contracts, respectively, where IPR contract is taken as the default category; P_{ijt} is the average ability of other workers in the same team as worker *i* captured by the method of Mas and Moretti (2009); R_t is a group formation dummy variable which takes one for self-selected team and zero for randomly-assigned team; S_{ijt} is a set of social preferences measured by artefactual field experiments; and X_{ijt} is a set of household's characteristics.

Our variables of interest is λ^{FW}, which captures the incentive effect of shifting from IPR to FW on individual productivity. If $\lambda^{FW} < 0$, we may conclude that there exists moral hazard induced by the lack of monetary incentives. In addition, a finding that λ^{GPR}<0 is consistent with the free-riding problem.

The estimated coefficients on (S_{ijt}, P_{ijt}, R_t), i.e., (δ, μ, π) capture the impacts of social preference, peer effects, and self-selection on worker incentives, respectively. Note that P_{ijt} is defined as the average of other worker's permanent ability rather than their contemporaneous productivity in order to avoid the reflection problem of Manski (1993, 2000). To weather the reflection problem, we adopt a two-step method developed by Mas and Moretti (2009).[20] As to the contract and team assignments, we select different wage contracts randomly and we assign self-selected teams or randomly assigned teams exogenously. Hence, by our field experiments, we believe OLS estimates of equation (1) can reveal the causal effects of contracts, peer effects, self-selection, and social preference

[20] In contrast, if we follow Hamilton, Nickerson, and Owan (2003) and use simple average productivity, the impact of average ability on individual productivity becomes uniformly larger. This indicates the problem of upward endogeneity bias arising from the reflection problem when we estimate peer effects.

on individual productivity fairly accurately.

4.2 The Ability Measure

The test of peer effects requires a measure of the permanent ability of worker *i* and her peers. To this purpose, we follow Mas and Moretti (2009) to estimate a worker productivity model with worker fixed effects considering the fact that an individual's productivity may be affected by co-worker composition.[21] Specifically, we employ the following estimation model:

$$l_{ijt} = \beta_i + TEAM_{jt}\gamma + \epsilon^a_{ijt} \quad (2)$$

where β_i is the individual fixed effects and $TEAM_{jt}$ are dummy variables for teams which are formed in each round. We construct each worker's permanent ability by estimating the worker fixed effects, β_i.

While, using these permanent abilities, we can construct other group members' average ability P_{ijt} and estimate Equation (1), we need to correct standard errors in this two-step procedure. Based on Mas and Moretti (2009), we employ the Bayesian parametric bootstrap method for this purpose: First, we estimate Equation (2) to elicit each worker's permanent ability. Using the estimated coefficients, we construct 1,000 simulated datasets by randomly drawing β_i from the distribution $N(\hat{\beta}, \hat{\Sigma})$ where $\hat{\beta}$ is the vector of the point estimates of β_i and $\hat{\Sigma}$ is the estimated variance-covariance matrix. Then, we estimate Equation (2) with these simulated datasets as well as the original dataset to obtain the coefficients for each datasets. The final standard errors are computed as $\sqrt{s_\beta^2 + \sigma_\beta^2}$, where s_β^2 is the sampling variance of the original datasets and σ_β^2 is the between-simulation variance which is the variance obtained from simulated datasets.

[21] Bandiera et al. (2010) adopt a similar strategy by limiting the sample to field days when the friends of each worker are not present.

4.3 Estimation Results

Table 3 shows that the regresssion results of estimating Equation (1) with full sample. In column (1), we estimate a simplified version of equation (1) without social preference variables, S_{ijt}. The result shows that IPR in the default category generated higher productivity than FW in which the average gap is around 3.6 meters per ten minutes. Yet, the difference between IPR and GPR is not statistically significant. The former finding is consistent with moral hazard problem, replicating the non-experimental results of Lazear (2000) and Foster and Rosenzweig (1994) and experimental results of Shearer (2004) which found the incentive effect generated by monetary incentives.

In column (2), we added a set of social preference variables elicited by artefactual field experiments, finding a significant role of non-monetary incentives. The offer amount difference between ultimatum and dictator games have positive and significant coefficients: people with guilt aversion preference are likely to exert more effort in rice planting works. Also, nonenviousness, which is definesd as the gap between the maximum value of the game, i.e. 10 coins, and the selected values of the minimum acceptance level as responder in ultimatum game, takes positive and significant coefficients. This suggests that nonenviousness stimulates positive work incentives in rice planting. Hence our results indicate that individual social preferences affect own productivity, underling intrinsic motivation.

In column (3) and (4), we add variables regarding to peer effects, i.e., average ability of other team members. The average ability of other team members are positive and significant, indicating postive spillover effects among workers in the workplace. Yet, the self-selected group formation variable has negative but insignificant effect. We can also closely look at the relationship variables: while common membership at the same church seems to be unrelated to productivity, kinship or close communal relationship with kabisiliya induces significantly negative effect on productivity unlike Bandiera et al. (2009). Finally, advanced payment has insignificant effect on effort, a finding which is not necessarily supporting the theoretical results by Braverman and Stiglitz (1982) that interlinked

contract of labor and credit transactions can mitigate moral hazard problems in labor contract.

Table 4 shows that estimation results of productivity equation (1) separately for each incentive schemes. Social preferences stimulate significantly posotive impacts only under FW according to both individual coefficient test and the joint significance test results. As for the individual coefficients, the sending amount in dictator game and the amount of offfer in ultimatum game, which cah be interpreted as degrees of altruism and fairness, respectively, are positively correlated with individual productivity under the FW contract. Also, the nonenviousness takes positive coefficients only for FW, indicating that nonenviousness excebates disincentive effects or moral hazard in the fixed wage contract. While in the case of FW, the joint F-test of the estimated coefficient of social preference variables reject the zero coefficients of these variables at 1%, the null hyopthesis cannot be rejected in IPR and GPR contracts. Also, the non-monetary peer effects captured directly by avarage ability of other team members have positive impact on individual producivity only in FW contract:

These results reconciles with the hypothesis that extrinsic incentives of individual piece rate would crowd out intrinsic motivatoin that persuade workers to restrain themselves from temptation of opportunistic behavior in the workplace (Bowles and Polania-Reyes, 2012).

Blinder-Oaxaca decomposition

To grasp the magnitude of producitivity effects of each component, we performed the Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973) for the average productivity difference between IPR and FW, which is 4.71 meters per ten minutes (Table 5). By construction of *quasi* within-subject design, the endowment difference is not statistically different from zero.[22] Using the Blinder-Oaxaca decomposition, we can investigate how much of this difference can be explained in terms of social preferences. Based on the estimation results of Table 4 column (4) and (5), the decomposition results are summarized in Table 5. As we can see, much of the performance difference between IPR and FW

[22] Since participating rice planters differ in each day, the endowment difference deviate from zero. Yet, if we employ balanced panel data, the endowment effect should be zero by construction.

can be decreased by the coefficient effects on social preference parameters. Comparing with IPR, FW seems to improve productivity by 2.94, 1.29, 4.75, and 2.91 meters through altruism, guilt aversion, nonenviousness, and voluntary contribution to public goods, respecitively.[23] This means that without social preference effects under FW, the performance gap between IPR and FW could have been more than tripled. In other words, introduction of monetary incentives under IPR crowded out more than half of motivation for work under IPR.

In sum, we obtain three main empirical findings. First, our basic results show the positive incentive effects in IPR, moral hazard problems in FW which is consistent with standard theoretical implications. Second, under FW, social preferences such as altruism, guilt aversion and nonenviousness play an important role in stimulating incentives but introducing monetary incentives crowds out such intrinsic motivations. Third, other non-monetary factors such as peer effects significantly change incentives under FW contract.

4.4 Testing Alternative Hypotheses

The Optimal Intensity of Incentive Condition

The results presented so far imply that the FW contract can exert effort by non-monetary incentives. An alternative hypothesis to show the optimality of FW is that principal's incapability to accurately measure worker's performance. We formulate a simple linearized model of the performance level of each worker observed by *kabisilya* as follows: $\eta = \delta e + u$, where e is a worker's true effort level; δ is *kabisilya*'s capacity to measure effort; and u is a measurement error. Then the rewards schedule offered by *kabisilya* is $w = F + \phi(\delta e + u)$, where ϕ is an intensity of incentives. The optimal choice of effort level for each worker is determined by incentive compatibility constraints:

[23] The estimated total social preference effect is 11.9 meters per 10 minutes.

$\delta\phi = C'(e)$ where $C(e)$ is an agent's convex effort cost function. In this setting, the sum of certainly equivalent for a principal and an agent is $P(e) - C(e) - \frac{1}{2}rC'(e)^2V$, where $P(e)$ is a principal's expected return function, r is coefficient of absolute risk aversion, and V is a variance of the effort measurement error, i.e., $V = var(u)$. Accordingly, the optimal intensity of incentives becomes (Milgrom and Roberts, 1992):

$$\phi = \frac{P'(e)\delta}{\delta^2 + rC"V} .$$

If $\delta = 0$, the optimal intensity of incentives ϕ becomes zero. This is the case where the FW contract becomes optimal. Hence, we can test the optimality of FW by testing a null hypothesis, $\delta = 0$, in the following estimable equation of the equation, $\eta = \delta e + u$:

$$\eta_i = \alpha_i + \delta l_{ijt} + X_{ijt}\gamma + u_{ijt}, \quad (3)$$

where we used our *observed* productivity data, l_{ijt}, as a proxy variable for the true effort. This implies that our estimates for δ involves attenuation bias which needs attentions. Yet, even in this case, we can still make an inference about the rejection of the null hypothesis, i.e., $\delta = 0$.

We use each *kabisilya*'s subjective evaluation regarding to each worker's performance as the dependent variable, η_i. We estimate the above equation for the FW and IPR samples separately and estimation results are shown in Table 6. In all specifications, outcomes observed by *kabisilya* are positively correlated with measured individual productivity in both fixed wage and individual piece rate. These results seem to be robust even if we include social connection variables to mitigate evaluation bias arising from favoritism. Moreover, considering possible errors in measuring the true effort level, l_{ijt}, the estimated δ is likely to involve attenuation bias. With these results and considerations, we can safely conclude that $\delta > 0$, rejecting FW as the optimal contract considering monetary incentives only. Hence, for FW to be optimal, it still seems necessary to incorporate non-monetary incentives.

Intertemporal Incentives

Another possibility for FW to be optimal is an intertemporal incentive in contract renewal. If contract renewal probability is a positive function of observed efforts, even under FW, expected wage level is contingent upon the observed effort level. In this case, FW contract can be *de facto i*ndividual piece rate contract. To test this hypothesis, we estimate the following model of contract renewals:

$$EXPER_i = \alpha_i + \rho\eta_{ij} + X_{ij}\gamma + u_{ij}$$

where $EXPER_i$ is years of planting experience under the current *kabisilya* as a proxy variable for contract renewals. The set of independent variables are the same as variables shown in equation (1). If the coefficient on the observed performance level, ρ, is positive and significant, then the model of intertemporal incentives will be supported. Table 7 shows estimation results. *Kabisilya*'s evaluation are only weakly related to years of planting under the current *kabisilya*. These estimation results are not necessarily consistent with the hypothesis of intertemporal incentives. Also, we can compare relative magnitude of social preference and intertemporal incentives using our empirical results. Using specification (4) in Table 4, we can quantify the impact of years of experiece on productivity: the point estimate of the variable is 0.075 (statisticaly insigificant) and the average experience length is 4.08 years, leading to producitivity gain of mere 0.31 meters evaluated at the average. Hence, the total magnitude of social preferences is much larger than that of intertemporal incentives.

5. Concluding Remarks

In this paper, we adopt a hybrid experimental method of framed and artefactual field experiments to examine the interplay between economic incentives and social norms. More specifically, in order to explain the seemingly inefficient FW contract in rice planting, we conduct

randomized control trials of three distinct labor contracts, i.e., FW, IPR, and GPR. Three main empirical findings emerge. First, our basic results show the positive incentive effects in IPR, moral hazard problems in FW, and weak free-riding behavior in GPR, which are consistent with implications of the standard agency theory well as empirical findings by Lazear (2000). Second, non-monetary incentives seem to play a significant role under FW. Under FW, an altruistic person with guilt averse and/or nonenvious preferences exert more work efforts in rice planting. Yet, these roles of social preferences disappear once we introduce monetary incentives of IPR contract. These results suggest that extrinsic incentives crowd out intrinsic motivation, a rare real world finding which confirms evidence from laboratory experiments reviewed by Bowles (2008). Also, other non-monetary factors such as peer effects significantly enhance work incentives under FW contract, which is consistent with findings by Mas and Moretti (2009). Finally, our results are not necessarily supportive to implications of the alternative hypotheses, i.e., interlinked contract of labor and credit transactions in mitigating moral hazard problems, the optimality of FW contract due to large effort measurement errors, and the existence of intertemporal incentives.

Our findings from a hybrid method of framed and artefactual field experiments imply that introduction of monetary incentives may generate inefficiency by crowding out intrinsic motivations. Hence, considering the interplay of intrinsic motivations and monetary incentives, this seemingly perverse traditional contractual arrangements might not be too inefficient. External validity of such a finding should be investigated rigorously by implementing carefully-designed filed experiments in other areas and industries in the future studies.

Reference

Bandiera, Oriana, Iwan Barankay, Imran Rasul (2005) “Social Preferences and the Response to Incentives: Evidence from Personnel Data” *Quarterly Journal of Economics*, Vol. 120, No. 3, pp. 917-962

Bandiera, Oriana, Iwan Barankay, Imran Rasul (2011) “Team Incentives: Evidence from a Firm Level Experiment”, *mimeographed.*

Bandiera, Oriana, Iwan Barankay, Imran Rasul (2010) “Social Incentives in the Workplace”, *Reivew of Economic Studies*, 77, 417-458.

Bénabou, Roland, and Jean Tirole (2003). “Intrinsic and Extrinsic Motivation.” *Review of Economic Studies*, 70(3): 489–520.

Blinder, Alan (1973), “Wage Discrimination: Reduced Form and Structural Estimates,” *Journal of Human Resources* 8, pp.436-455.

Bowles, Samuel (2008), “Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments,” *Science* 320 (5883), pp. 1605-1609.

Bowles, Samuel and Sandra Polanía-Reyes (2012) “Economic Incentives and Social Preferences: Substitutes or Complements?,” *Journal of Economic Literature*, 50:2, 368–425

Braverman, Avishay and Joseph E. Stiglitz (1982), "Sharecropping and the Interlinking of Agrarian Markets," *American Economic Review* 72 (4), 695-715.

Bruce Shearer (2004) “Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment.” *Review of Economic Studies*, 71 (2), 513-534.

Camerer, Colin F. (2003), *Behavioral Game Theory*: *Experiments in Strategic Interaction*, Princeton University Press.

Camerer, Colin F., and Ernst Fehr (2004) “Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists”, in Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin F. Camerer, Ernst Fehr, and Herbert Gintis (eds.) *Foundations of Human Sociality:*

Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies, Oxford: Oxford University Press, pp. 55–95.

Cardenas, Juan Camilo and Jeffry Carpenter (2008) "Behavioural Development Economics: Lessons from Field Labs in the Developing World" *Journal of Development Studies* 44(3), pp311-338

Carpenter, Jeffrey, Samuel Bowles, Herbert Gintis, Sung-Ha Hwang (2009) "Strong reciprocity and team production: Theory and evidence" *Journal of Economic Behavior & Organization* 71 pp. 221–232

Carpenter, Jeffrey and Erika Seki (2010) "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay" *Economic Inquiry*, Volume 49, Issue 2, pp. 612–630

Carpenter, Jeffrey and Tyler Williams (2014) "Peer Monitoring and Microcredit: Field Experimental Evidence from Paraguay," *Oxford Development Studies,* Vol. 42, No. 1, 111–135.

David, Cristina C., Violeta G. Cordova, and Keijiro Otsuka (1994), "Technological Change, Land Reform, and Income Distribution in the Philippines," in David, C.C. and Otsuka, K. (1994). *Modern Rice Technology and Income Distribution in Asia*. Lynne Rienner Publishers, Boulder, Colorado.

Deninger, Klaus., P. Olinto, and M. Maertens (2000) "Redistribution, investment, and human capital accumulation: the case of agrarian reform in the Philippines,"*mimeographed*, World Bank.

Eriksson, Tor and Marie Claire Villeval (2008) "Performance-pay, sorting and social motivation" *Journal of Economic Behavior & Organization* 68: 412–421

Estudillo, Jonna P., Yasuyuki Sawada, Mahabub Hossain (2005) "Socioeconomic Changes and the Decline in Poverty: A View from Three Villages in the Philippines, 1985–1997", *Asian Economic Journal* 2005, Vol. 19 No. 4, pp.383-406

Estudillo, Jonna P., Yasuyuki Sawada, Keijiro Otsuka (2006) "The Green Revolution, development of labor markets, and poverty reduction in the rural Philippines, 1985–2004", *Agricultural Economics* 35 pp.399–407

Estudillo, Jonna P., Yasuyuki Sawada, Keijiro Otsuka (2008) “Poverty and Income Dynamics in Philippine Villages, 1985–2004”, *Review of Development Economics*, Vol. 12(4), pp. 877-890

Estudillo, Jonna P., Yasuyuki Sawada, Keijiro Otsuka (2009) “The Changing Determinants of Schooling Investments: Evidence from Villages in the Philippines, 1985–89 and 2002–04”, *Journal of Development Studies*, Vol. 45, No. 3, pp.391–411

Falk, Armin and Andrea Ichino (2006) “Clean Evidence on Peer Effects” *Journal of Labor Economics*, Vol. 24, No. 1, pp. 39-57

Fehr, Ernst and Klaus M. Schmidt (1999) “A Theory of Fairness, Competition, and Cooperation” *Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-868

Foster, Andrew D and Mark R. Rosenzweig (1994) "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health," *Review of Economics and Statistics* vol. 76(2), 213-27.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan. (2003) “Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation.” *Journal of political Economy*, 111(3): 465–97.

Harrison, Glenn W. and John A. List (2004), “Field Experiments,” *Journal of Economic Literature* 42(4), 1009-1055.

Hart, Oliver and Holmstrom, Bengt (1987) “The theory of contracts”, in *Advances in Economic Theory*, T. Bewley ed., Cambridge University Press.

Hayami, Yujiro and Masao Kikuchi (1978) *Anatomy of a Peasant Economy: A rice village in the Philippines*, The International Rice Research Institute

Hayami, Yujiro and Masao Kikuchi (2000) *A Rice Village Saga: Three Decades of Green Revolution in the Philippines*, Lanham/Boulder/New York: Barnes & Noble in association with IRRI

Itoh, Hideshi (2004), “Moral Hazard and Other-Regarding Preferences,” *Japanese Economic Review* 55(1), 18-45.

Kandel, Eugene and Edward P. Lazear (1992) “Peer Pressure and Partnerships”, *Journal of Political*

Economy, 100, 801-813.

Kerkvliet, Benedict J. Tria. (1990) *Everyday Politics in the Philippines: Class and Status Relations in a Central Luzon Village*. Berkeley: University of California Press.

Lazear, Edward P. (1995) *Personnel Economics*, MIT Press. Cambridge

Lazear, Edward P. (2000) "Performance pay and productivity" *American Economic Review*, Vol. 90, No. 5, pp.1346-1361

Levitt, Steven D. and John A. List (2009), "Field Experiments in economics: The past, the present, and the future," *European Economic Review* 53, 1–18

Levitt, Steven D. and John A. List (2007) "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?" *Journal of Economic Perspectives* 21(2), pp. 153-174

Ledyard, John (1995) "Public Goods: A Survey of Experimental Research," in *Handbook of Experimental Economics*, J. Kagel and A. Roth (Eds.), pp. 111–194. Princeton, NJ: Princeton University Press

List, John and Imran Rasul (2011), "Field Experiments in Labor Economics," in *Handbook of Labor Economics* Vol.

Mas, Alexandre, Enrico Moretti (2009) "Peers at Work", *American Economic Review*, 99, 112-145.

Munro, Alistair, Arjan Verschoor, and Amaresh Dubey (2013) "Does working with spouses make teams more productive? A field experiment in India using NREGA," *Economic Letters* 118, pp.506-508.

Oaxaca, Ronald (1973), "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14: 693-709.

Otsuka, Keijiro (1991) "Determinants and consequences of land reform implementation in the Philippines," *Journal of Development Economics* 35(2): pp.339-355.

Otsuka, Keijiro, Hiroyuki Chuma, Yujiro Hayami (1993) "Permanent Labour and Land Tenancy Contracts in Agrarian Economies: An Integrated Analysis", *Economica* 60, pp.57-77

Radner, Roy (1986) "Repeated Partnership Games with Imperfect Monitoring and No Discounting" *Review of Economic Studies*, Vol. 53, No. 1, pp. 43-57.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir (1991), "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study," *American Economic Review* Volume 81, Issue 5 , 1068-1095.

Roumasset, James and Lee, Sang-Hyop Lee (2007) "Labor: Decisions, Contracts and Organization," Robert Evenson & Prabhu Pingali, eds., *Handbook of Agricultural Economics*, Elsevier, chapter 52, 2705-2740.

Sausgruber, Rupert (2009) "A note on peer effects between teams" *Experimental Economics* 12: 193–201

Schechter, Laura (2007), "Traditional trust measurement and the risk confound: An experiment in rural Paraguay," *Journal of Economic Behavior & Organization* Vol. 62, 272–292.

Shearer, Bruce (2004) "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment" *Review of Economic Studies,* Vol. 71, No. 2, pp. 513-534.

Umekage, Osamu, Koichiro Inoue, and Tadayo Watabe (1966), "Studies on the Simplification of Transplanting Method of Rice Plants I: Growth Habit of the Plants of Which Seedlings are Horizontally Transplanted (in Japanese)," *Kyoto Prefectural University Academic Report on Agriculture* (Kyoto Furitsu Daigaku Gakujyutsu Houkoku Nougaku) 18, 70-77.

Figure 1: Structure of a planting group

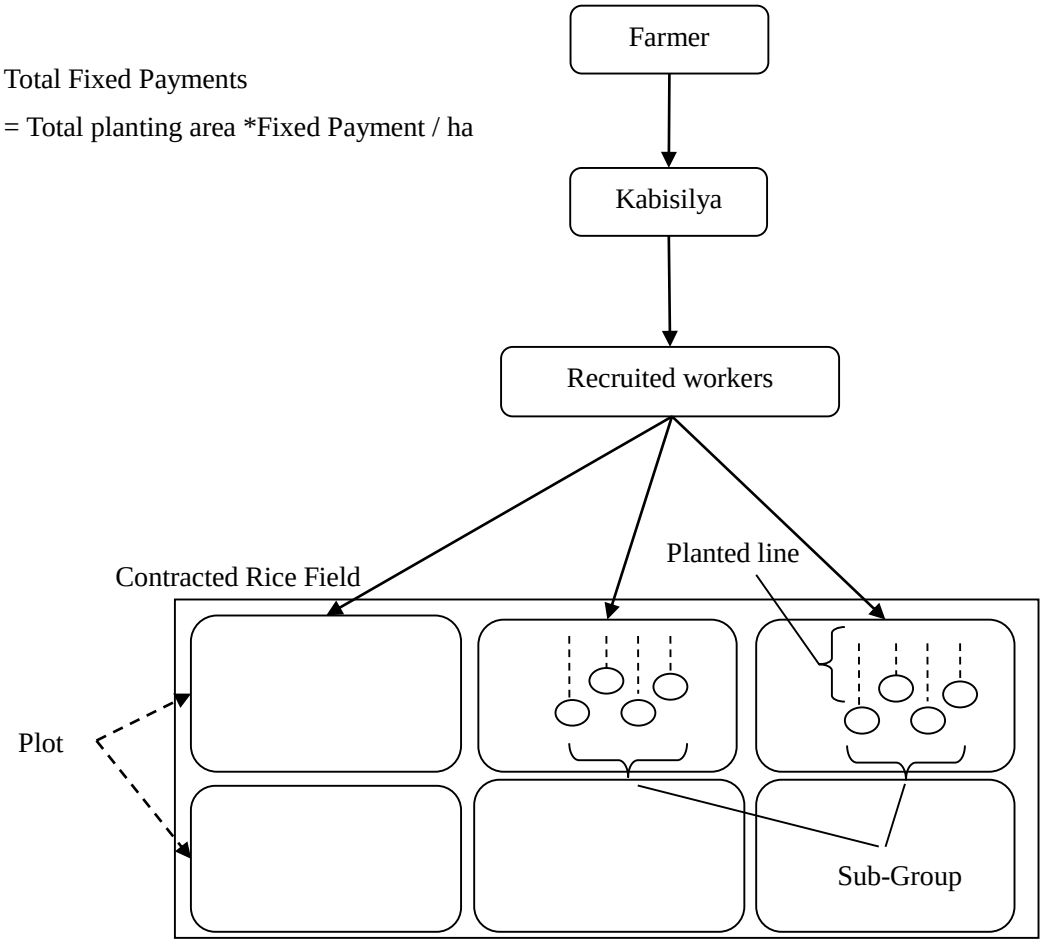

Figure 2: Timeline of experiments

Date (Jan, 2011)		11		12		13		14		15		17		18		19		20		21		22		24		
Time		AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM	PM	AM		
Selection		Random-selection								Self-selection								Artefactual field experiment		Self-selection						
G village	Round	1	2	3		4	5	6	7	8	9	10				11	12	13				14	15		Sub-total	Total
	FW	x				x			x			x					x						x		6	15
	GPR			x				x		x						x									4	
	IPR		x				x				x							x				x			5	

Selection		Random-selection									Self-selection						Artefactual field experiment	Self-selection			Sub-total	Total
M village	Round	1		2	3		4				5	6	7	8	9	10		11	12	13		
	FW	x					x					x			x				x		5	13
	GPR				x								x	x				x			4	
	IPR			x							x					x				x	4	

Note) x indicates the experiment we conducted; and * the first two rounds are dropped from our analysis because the fixed wage component is set at 20 PHP in these rounds.

Table 1: Descriptive statistics

Panel A; Productivity	Obs	Mean	Std. Dev.	Min	Max
Productivity (stack for 1-10, 10-20, 20-30min)					
Fixed wage	732	26.07632	9.226555	1.896	69.144
Individual piece rate	522	29.96725	9.370332	2.01915	65.388
Group piece rate	453	29.74732	9.205822	6.4725	69.8
Panel B: Filed Experiment Condition	Obs	Mean	Std. Dev.	Min	Max
Average years of aquintance with group members	569	10.63813	9.437524	0	42.4
Share of group members attending to the same	569	0.3928237	0.3917962	0	1
Self-selection dummy	569	0.4604569	0.4988725	0	1
Plot condition (base category: bad)					
Usual	569	0.5465729	0.4982643	0	1
Good	569	0.3585237	0.4799889	0	1
Weather (base category: clear and sunny)					
Sunny	569	0.3216169	0.4675078	0	1
Cloudy	569	0.2565905	0.4371357	0	1
Wind force (base category: no wind)					
Breeze	569	0.4973638	0.500433	0	1
Weak	569	0.059754	0.237239	0	1
Strong	569	0.1757469	0.3809396	0	1
Missing wind force	569	0.0228471	0.1495474	0	1
Temperature	569	22.93146	12.99501	0	38
Missing temperature	569	0.2337434	0.4235832	0	1
Health condition (base category: good)					
Same as usual	548	0.4744526	0.4998031	0	1
Bad	548	0.0145985	0.1200488	0	1
Experimenter dummy	569	0.4253076	0.4948246	0	1
G village dummy	569	0.5202109	0.5000309	0	1
Plant consecutively	569	0.7486819	0.4341527	0	1
PM session dummy	569	0.4956063	0.5004206	0	1
Interval (base category: 0-10min)					
10-20min	569	0.340949	0.4744454	0	1
20-30min	569	0.3268893	0.4694891	0	1
Group size	569	5.369069	0.8417951	3	7
Participation number	569	3.620387	2.228182	1	10
Panel C: Individual Characteristics	Obs	Mean	Std. Dev.	Min	Max
# of participation / total field experiment sessions	115	0.4138036	0.2122222	0.0769231	0.8181818
Occupation type (base category: non-agricultural)					
Farmer	110	0.2	0.4018307	0	1
Daily agricultural worker (landless)	110	0.4181818	0.4955179	0	1
Porcientuhan	110	0.2909091	0.4562603	0	1
Other	110	0.0090909	0.0953463	0	1
Age	111	35.59459	15.89304	13	80
Age (squared)	111	1517.288	1257.784	169	6400
Sex (male=1)	111	0.3423423	0.4766454	0	1
Years of schooling	109	7.036697	2.613805	1	13
Years of experience under current kabisiliya	104	4.076923	5.354324	0	30
Advanced payment (PHP)	102	625.4902	617.1936	0	3000
Relationship with kabisiliya					
Relative or family of kabisiliya	104	0.3846154	0.4888602	0	1
Friend of kabisiliya	104	0.2788462	0.4506033	0	1
Neighbor of kabisiliya	104	0.1826923	0.3882853	0	1

Figure 3: Cumlative distribution function of individual productivity

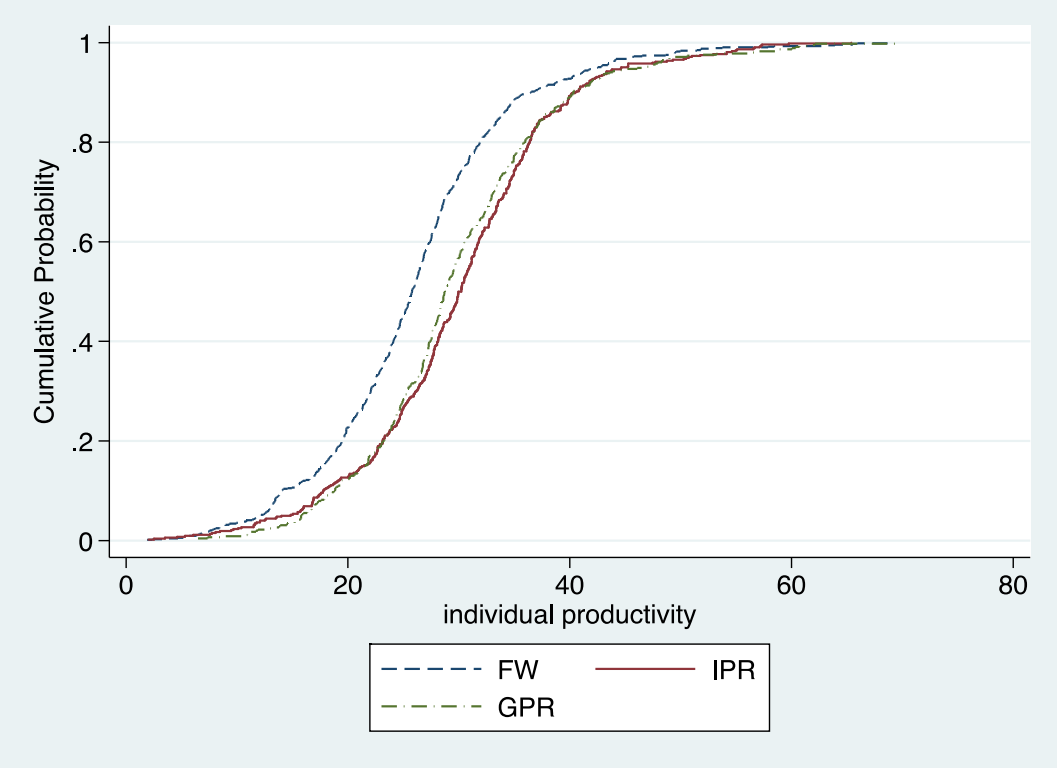

Figure 4: Individual and average productivity of other peers

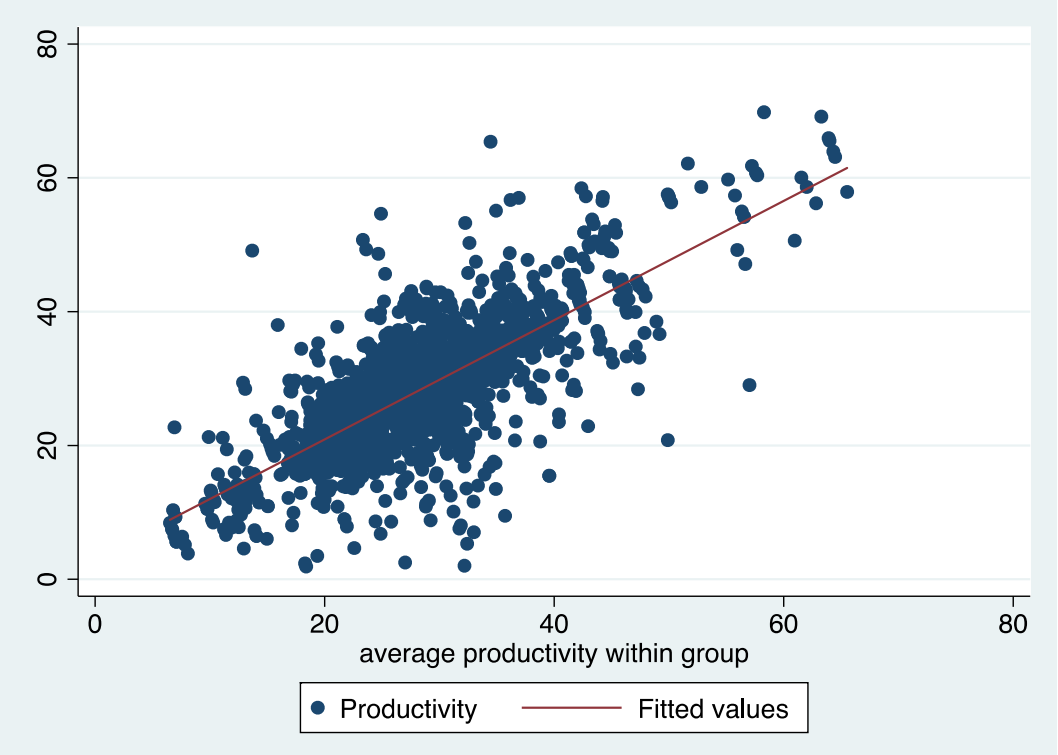

Table 2: Descriptive statistics of artefactual field experiments

Panel A: Public Goods Game	Obs	Mean	Std. Dev.	Min	Max
Contribution in 1st round	104	33.26923	20.64223	0	100
Contribution in 2nd round	104	31.44231	21.8298	0	100
Contribution in 3rd round	104	35.19231	27.51807	0	100
Average contribution	104	33.30128	19.21038	10	86.66666
Monitoring in 1st round	104	0.4615385	0.5009327	0	1
Monitoring in 2nd round	104	0.3461538	0.4780468	0	1
Monitoring in 3rd round	104	0.3076923	0.4637735	0	1
# of disapproval messages in 1st round	104	0.1346154	0.3955503	0	2
# of disapproval messages in 2nd round	104	0.2115385	0.5689895	0	3

Panel B: Dictator and Ultimatum Game	Obs	Mean	Std. Dev.	Min	Max
Dictator game (sending amount)	104	27.21154	18.45786	0	100
Ultimatum game (sending amount)	103	35.04854	15.77251	10	90
Ultimatum game (responder's choice)	104	2.865385	2.755604	0	10

Panel C: Risk Game	Obs	Mean	Std. Dev.	Min	Max
Invested amount	104	37.69231	22.4357	10	100

Figure 5: Mean contribution level in public goods game

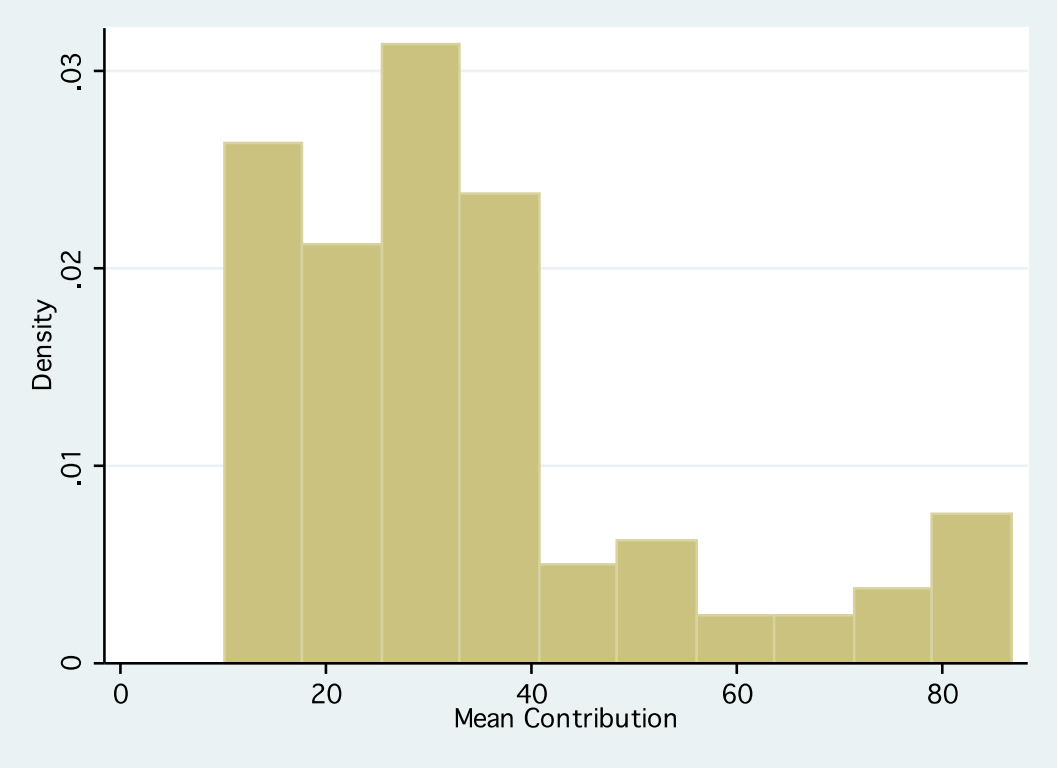

Figure 6: Proposer's offer in dictator game

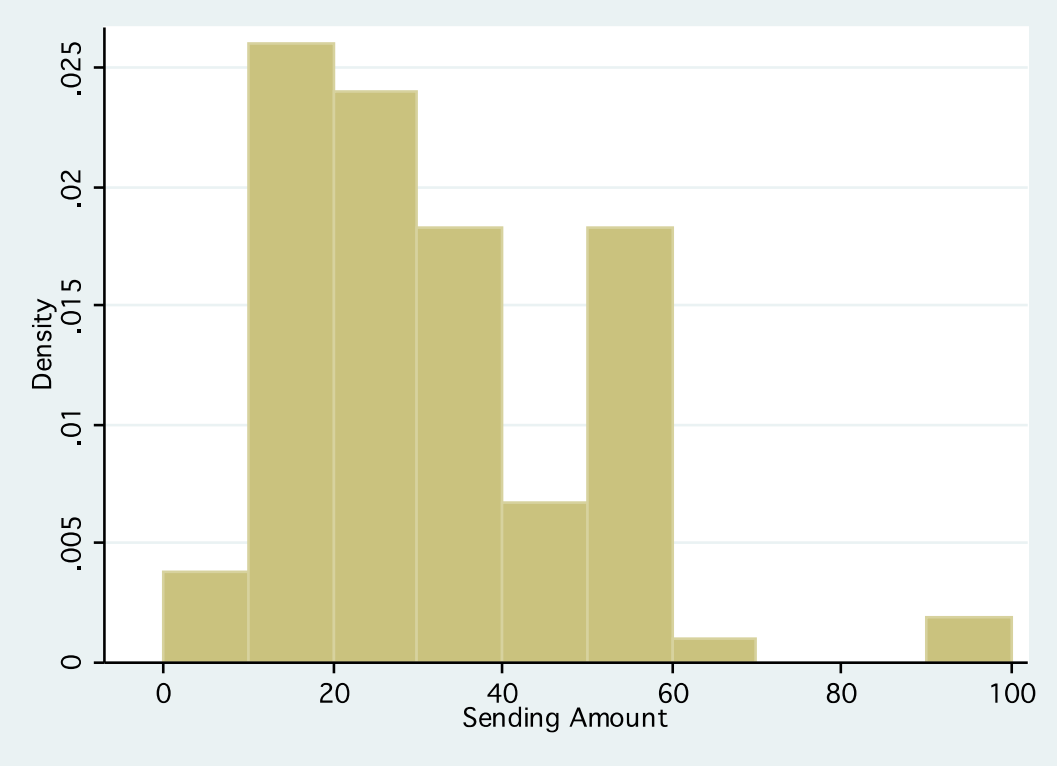

Figure 7: Proposer's offer in ultimatum game

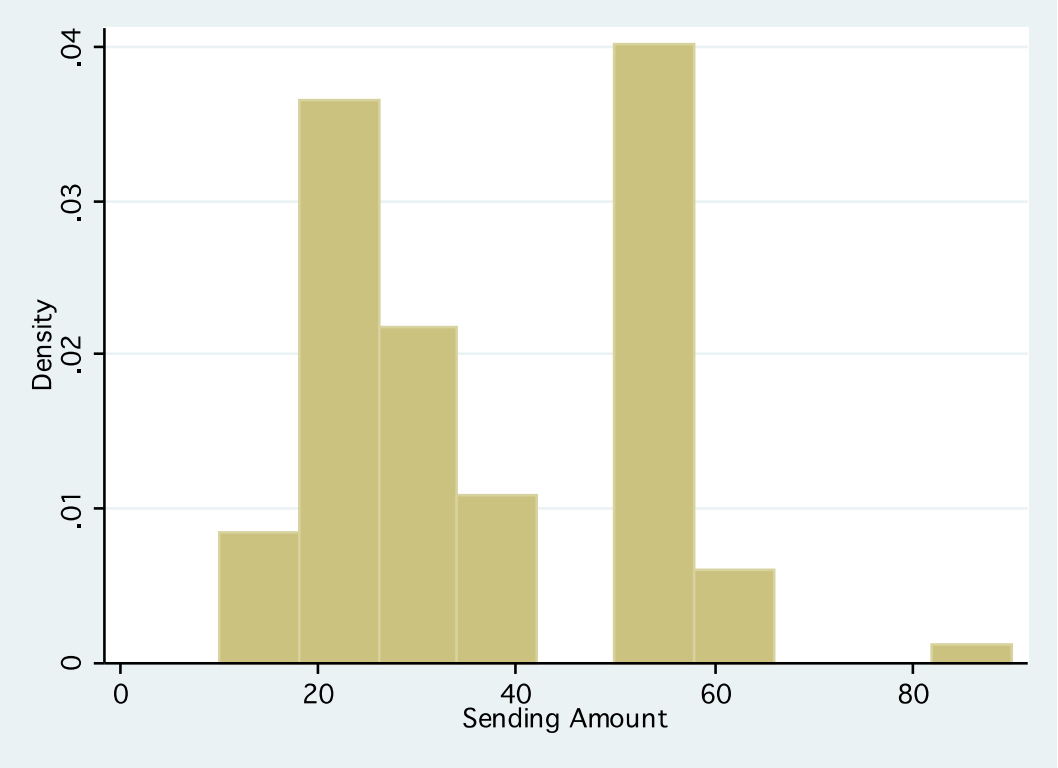

Figure 8: Responder's acceptance in ultimatum game

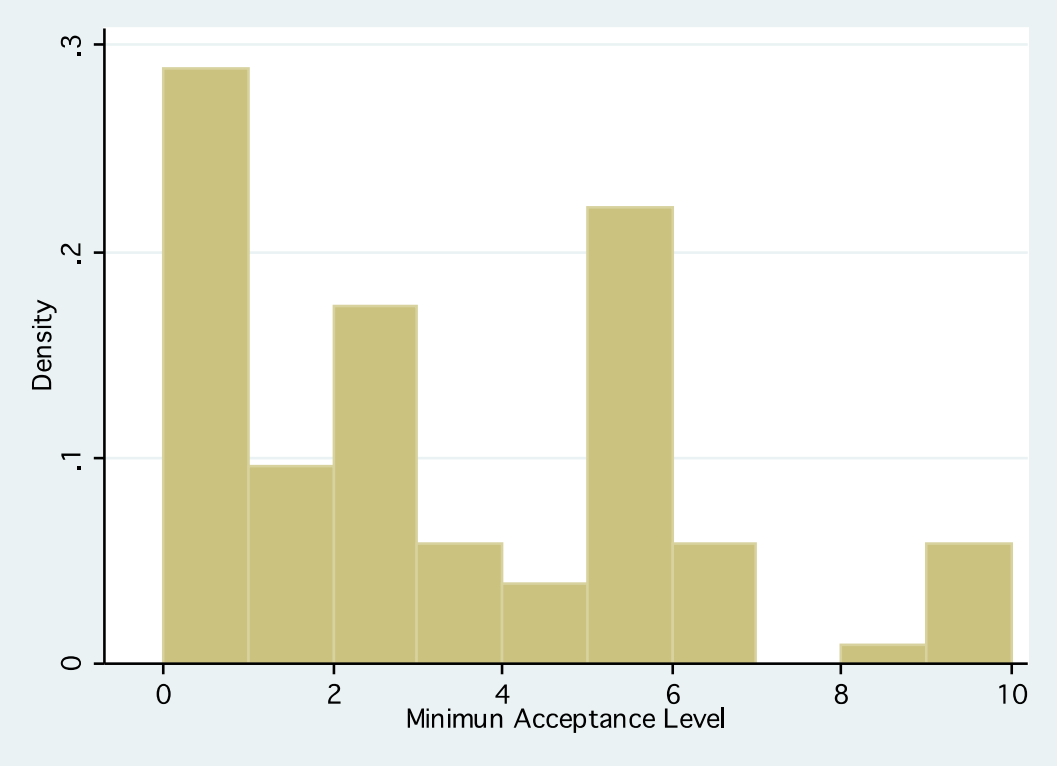

Figure 9: Investment level in risk game

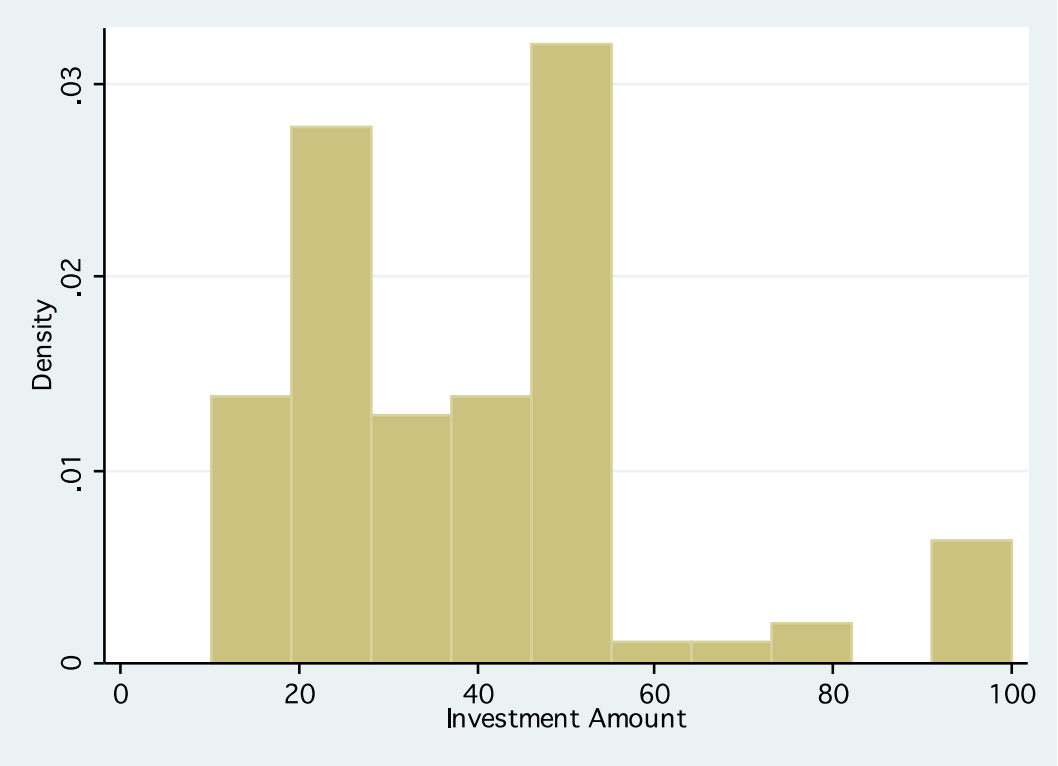

Figure 10 Distribution of estimated permanent ability

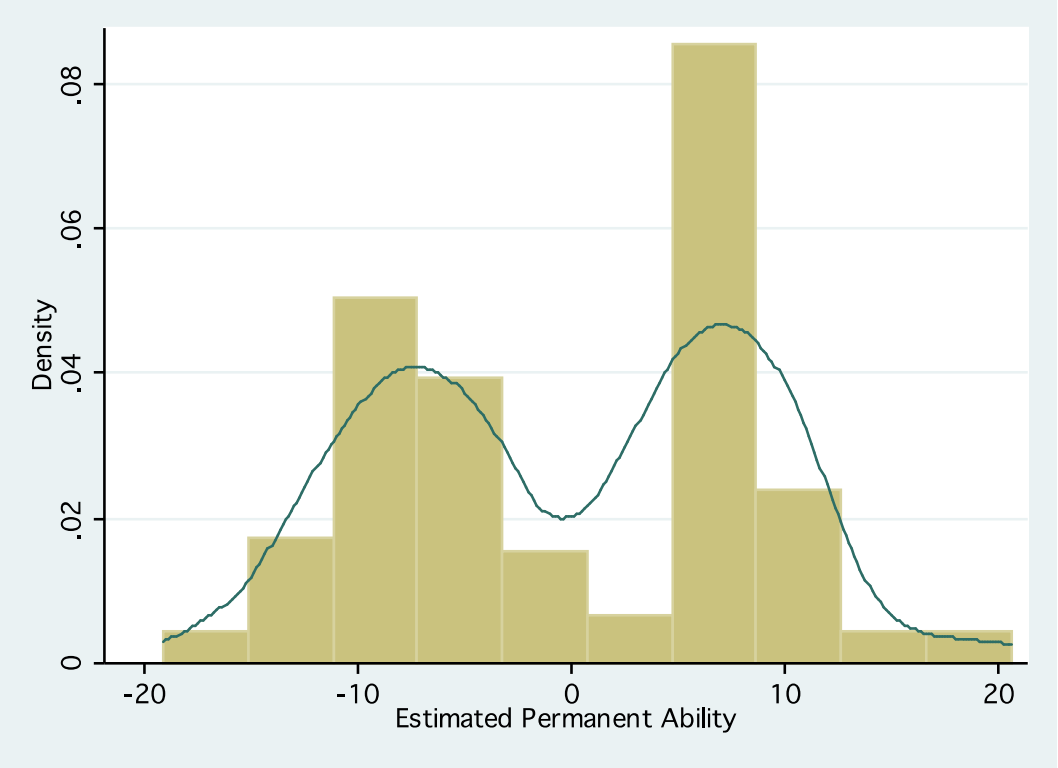

Table 3: Estimation results of rice planter's work performance regression

(Dependent variable is the length of planted line per ten minutes)

	(1)	(2)	(3)	(4)
Fixed wage dummy	-3.608***	-3.651***	-3.665***	-3.745***
	(1.022)	(1.074)	(0.870)	(0.868)
Group piece rate dummy	-0.808	-1.266	-0.963	-0.870
	(1.016)	(1.030)	(0.868)	(0.867)
Dictator game		0.0531	0.0796***	0.0822***
		(0.0328)	(0.0222)	(0.0229)
Guilt aversion (amount sent in ultimatum game		0.0905***	0.0918***	0.0775***
- amount sent in dictator game)		(0.0268)	(0.0212)	(0.0231)
Nonenviousness (10 coins – minimum acceptance		0.465***	0.402***	0.463***
Level in ultimatum game)		(0.172)	(0.108)	(0.111)
Average cooperation in PGG		0.0257	0.00751	-0.00225
		(0.0266)	(0.0203)	(0.0214)
Risk game		-0.0277	-0.0406***	-0.0389**
		(0.0203)	(0.0157)	(0.0164)
Average ability of other group members			0.749***	0.742***
			(0.184)	(0.184)
Self-selection dummy			-0.692	-0.869
			(1.026)	(1.033)
Years of experience under current kabisiliya			0.0732	0.0534
			(0.0554)	(0.0570)
Relative or family of kabisiliya				-2.403***
				(0.924)
Friend of kabisiliya				-2.031**
				(0.918)
Neighbor of kabisiliya				-0.141
				(1.122)
Average years of aquintance with group members				0.0597*
				(0.0335)
Share of group members attending to the same church				0.888
				(0.723)
Advanced payment (PHP)			0.000707	0.000251
			(0.000451)	(0.000468)
# of participation / total field experiment sessions	8.853***	8.611***	5.903***	6.744***
	(2.949)	(2.923)	(2.119)	(2.170)
Field Experiment Condition	YES	YES	YES	YES
Individual Characteristics	YES	YES	YES	YES
Observations	1584	1488	1371	1371
R-squared	0.186	0.206	0.230	0.239

Clustered SEs are reported in parentheses in (1) and (2), and bootstrapped SEs in (3) and (4).

Field Experiment Condition: plot condition dummies, weather dummies, wind force dummies and its missing dummy, temperature and its missing dummy, health condition dummies, G village dummy, experimenter dummy, consecutive planting dummy, PM dummy, interval dummies, group size dummies, participation number dummies

Individual Characteristics: occupation type dummies, age, age squared, sex, years of schooling

Estimation results of rice planter's work performance regression by contract type
(Dependent variable is the length of planted line per ten minutes)

	(1)	(2)	(3)	(4)	(5)	(6)
Type of contract	FW	IPR	GPR	FW	IPR	GPR
Dictator game	0.0925*	-0.0346	0.0283	0.119*	0.00532	0.0448
	(0.0302	(0.039)	(0.037)	(0.0333	(0.0492)	(0.0446)
Guilt aversion (amount sent in ultimatum game	0.106**	-0.00152	0.0854*	0.0978	-0.0575	0.108**
- amount sent in dictator game)	(0.0310	(0.0407)	(0.0369)	(0.0349	(0.0514)	(0.0478)
Nonenviousness (10 coins – minimum acceptance	0.614**	-0.153	0.403**	0.585*	-0.146	0.443**
Level in ultimatum game)	(0.184)	(0.170)	(0.199)	(0.190)	(0.183)	(0.219)
Average cooperation in PGG	0.0333	-0.0418	0.0347	0.0204	-0.0773*	0.0298
	(0.0268	(0.0382)	(0.0348)	(0.0305	(0.0450)	(0.0419)
Risk game	-0.0598***	0.0399	-0.0503*	-	-0.00735	-
	(0.0227)	(0.0280)	(0.0262)	(0.0238	(0.0333)	(0.032
Average ability of other group members	0.730**	0.168	0.464	0.666*	0.290	0.668
	(0.266)	(0.400)	(0.379)	(0.277)	(0.442)	(0.410)
Self-selection dummy				-11.16	-4.035	-
				(9.177)	(3.052)	(2.041)
Years of experience under current kabisiliya				0.0745	0.237**	-0.126
				(0.0859	(0.109)	(0.110)
Relative or family of kabisiliya				-2.288*	-0.887	-3.694**
				(1.360)	(1.804)	(1.734)
Friend of kabisiliya				-1.422	0.851	-4.301**
				(1.329)	(1.861)	(1.753)
Neighbor of kabisiliya				1.270	0.243	-4.127*
				(1.734)	(2.037)	(2.293)
Average years of acquaintance with group members				0.0997	-0.0447	0.0303
				(0.0519	(0.0679)	(0.0605)
Share of group members attending to the same church				1.491	2.839*	0.148
				(1.075)	(1.487)	(1.466)
Advanced payment (PHP)				-3.59e-	0.00025	0.000323
				(0.0007	(0.0008)	(0.0010)
# of participation / total field experiment sessions	4.550	16.39**	3.264	3.310	16.63**	0.890
	(3.559)	(4.107)	(3.655)	(3.646)	(4.607)	(4.455)
Field Experiment Condition	YES	YES	YES	YES	YES	YES
Individual Characteristics	YES	YES	YES	YES	YES	YES
F stat on social preference	7.84***	0.98	1.18	11.25*	1.50	1.23
Observations	627	456	405	582	426	363
R-squared	0.268	0.312	0.348	0.314	0.296	0.374

Bootstrapped SEs are reported in parentheses.
F-stat on social preferences are calculated based on clustered standard errors.
Field Experiment Condition: plot condition dummies, weather dummies, wind force dummies and its missing dummy, temperature and its missing dummy, health condition dummies, G village dummy, experimenter dummy, consecutive planting dummy, PM dummy, interval dummies, group size dummies, participation number dummies
Individual Characteristics: occupation type dummies, age, age squared, sex, years of schooling

Table 5: Blinder-Oaxaca Decomposition of the Performance Difference between Fixed Wage and Individual Piece Rate

VARIABLES	(1) Differential	(2) Endowments	(3) Coefficients	(4) Interaction
Prediction (Fixed wage)	25.77***			
	(0.542)			
Prediction (Individual piece rate)	30.48***			
	(0.578)			
Difference	-4.711***			
	(0.711)			
Dictator game		0.00802	2.944**	0.171
		(0.0780)	(1.300)	(0.213)
Guilt aversion (amount sent in ultimatum game		-0.0142	1.291**	0.0383
- amount sent in dictator game)		(0.102)	(0.561)	(0.275)
Nonenviousness (10 coins – minimum acceptance		-0.0929	4.754***	0.466**
Level in ultimatum game)		(0.122)	(1.374)	(0.232)
Average cooperation in PGG		-0.292	2.906*	0.369
		(0.211)	(1.605)	(0.249)
Risk game		-0.0289	-1.494	-0.172
		(0.104)	(1.355)	(0.170)
Average ability of other group members		0.548	-0.882	0.709
		(0.801)	(1.228)	(0.965)
Self-selection dummy		-0.0797	-3.310	-0.141
		(0.142)	(4.157)	(0.295)
Years of experience under current kabisiliya		-0.0879	-0.745	0.0604
		(0.128)	(0.494)	(0.0933)
Relative or family of kabisiliya		-0.0375	-0.641	-0.0592
		(0.0684)	(0.840)	(0.0964)
Friend of kabisiliya		-0.0100	-0.624	0.0267
		(0.0372)	(0.541)	(0.0867)
Neighbor of kabisiliya		-0.00838	0.152	-0.0354
		(0.0587)	(0.335)	(0.0818)
Average years of acquaintance with group members		0.0242	1.692*	-0.0783
		(0.0521)	(0.937)	(0.134)
Share of group members attending to the same church		0.143	-0.531	-0.0680
		(0.115)	(0.635)	(0.0927)
Advanced payment (PHP)		-0.0145	-0.220	0.0165
		(0.0472)	(0.679)	(0.0532)
# of participation / total field experiment sessions		-0.496	-7.344***	0.398
		(0.321)	(2.405)	(0.268)
Total		1.700	-7.216	0.804
		(1.330)	(6.196)	(6.260)
Field Experiment Condition	YES	YES	YES	YES
Individual Characteristics	YES	YES	YES	YES
Observations	1008	1008	1008	1008

Bootstrapped SEs are reported in parentheses.
Field Experiment Condition: plot condition dummies, weather dummies, wind force dummies and its missing dummy, temperature and its missing dummy, health condition dummies, G village dummy, experimenter dummy, consecutive planting dummy, PM dummy, interval dummies, group size dummies, participation number dummies
Individual Characteristics: occupation type dummies, age, age squared, sex, years of schooling

Table 6: Test of optimal intensity of incentives

(Dependent variable is the kabisiliya's subjective evaluation of each planting worker's performance)

	(1) FW	(2) FW	(3) FW	(4) IPR	(5) IPR	(6) IPR
Productivity	0.0437***	0.0350**	0.0357**	0.0769***	0.0451***	0.0457***
	(0.0163)	(0.0149)	(0.0147)	(0.0170)	(0.0155)	(0.0150)
Relative or family of kabisiliya			0.462			1.136
			(0.769)			(1.055)
Friend of kabisiliya			0.456			0.766
			(0.850)			(1.099)
Neighbor of kabisiliya			0.304			1.000
			(0.864)			(0.933)
# of participation / total field experiment sessions		1.544	1.414		1.328	1.090
		(2.043)	(2.069)		(2.134)	(2.255)
Field Experiment Condition	NO	YES	YES	NO	YES	YES
Individual Characteristics	NO	YES	YES	NO	YES	YES
Observations	678	627	627	492	459	456
R-squared	0.023	0.391	0.394	0.066	0.514	0.526

Clustered SEs are in parenthese.

Field Experiment Condition: plot condition dummies, weather dummies, wind force dummies and its missing dummy, temperature and its missing dummy, health condition dummies, G village dummy, experimenter dummy, consecutive planting dummy, PM dummy, interval dummies, group size dummies, participation number dummies

Individual Characteristics: occupation type dummies, age, age squared, sex, years of schooling

Table 7: Test of intertemporal incentives

(Dependent variable is years of work experience under the current kabisiliya)

VARIABLES	(1)	(2)	(3)
Kabisiliya's subjective evaluation	0.308*	0.305	0.310
	(0.166)	(0.207)	(0.215)
Relative or family of kabisiliya			-1.294
			(1.960)
Friend of kabisiliya			0.774
			(2.259)
Neighbor of kabisiliya			-2.211
			(1.781)
Individual characteristics	NO	YES	YES
Observations	103	100	100
R-squared	0.024	0.262	0.300

Clustered SEs are in parentheses. Individual Characteristics: # of participation / total field experiment sessions, occupation type dummies, age, age squared, sex, years of schooling, G village dummy