Incentives, Self-selection, and Social Norms in the Labor Contract: A Two-stage Field Experiment in the Philippines

Jun Goto* Yuki Higuchi† Takeshi Aida‡ Yasuyuki Sawada§

August 31, 2018

Abstract

This paper decomposes productivity difference between fixed wage (FW) contract and individual piece rate (IPR) contract into self-selection and incentive effects, using a unique two-stage field experiment. We offered an option of switching to IPR contract for agricultural workers in the Philippines, whose default option has traditionally been FW contract, and we converted random half of those who opted for IPR contract back to the original FW contract. By comparing three groups, i.e., those who chose and worked under IPR contract, those who chose IPR but worked under FW contract, and those who chose and worked under FW contract, we find that the self-selection effect accounts for 60% of the productivity difference between the two types of contract. By combining with the data collected from lab-in-the field experiments, we find that the choice of IPR contract is associated with social norm parameters, namely, inequity aversion and kinship tax rate. Exploiting our random group assignment, we also find that the influence of social norm is particularly strong when a worker anticipates high probability of repeated interaction with other group members.

Keywords: social norm, inequity aversion, kinship tax, community enforcement

*Kobe University; email: gotojun.jg@gmail.com

†Nagoya City University; email: higuchi@econ.nagoya-cu.ac.jp

‡Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO); email: aidatakeshi@gmail.com

§Asian Development Bank; email: ysawada@adb.org

1 Introduction

One of the most important tenets in economics is among other things that individual behavior is influenced by both economic incentives and social norms, which was originally put forward by Adam Smith (1759; 1776). Economic incentives in labor contracts such as a bonus payment proportional to individual outcomes work through two mechanisms: first, individuals positively respond to the intensity of incentives (incentive effect); second, individuals who are more capable and perform better than others sort themselves into a contract with the higher intensity of incentives out of various options (self-selection effect). Disentangling these two effects is indispensable since each of them has different implications for how to improve efficiency of labor markets as a whole: an incentive effect implies that economic incentives have a (positive) causal impact on worker's effort level whereas a self-selection effect doesn't indicate that an incentive prompts workers to exert more effort and can simply be interpreted as labor reallocation between available contract options. Therefore, introduction of an economic incentive may not have any impact on efficiency if just a self-selection effect operates and a worker's mobility between contracts is not allowed.

Obviously, economic contracts are embedded in complex social contexts. Then, as mentioned by Adam Smith, social norms can shape above two effects. In fact, recent studies corroborate the impacts of social consideration on incentive effects: social preferences and social norms can prominently alter the way that people react to economic incentives (Bandiera, Barankay, & Rasul, 2005, 2009, 2010; Goto, Sawada, Aida, & Aoyagi, 2015). Few studies, however, unveil how these social incentives shape *self-selection processes* despite the fact that understanding their mechanisms are of importance to design optimal contracts in labor markets and to settle the policy debate on whether introducing variable payment schemes results in an increase of firm's productivity as a whole.

Dohmen and Falk (2011) is the exceptional and seminal work in providing the first experimental evidence on the roles of social preferences in sorting processes. They conduct a controlled laboratory experiment in which subjects are asked to select one of the contracts between the fixed payment and variable payments such as individual piece rate, revenue-sharing, and tournament. They find that productivity gap between fixed and variable payments is largely driven by a selection process and individuals' trust is positively correlated with the choice of individual piece rate. Their study provokes several open questions which are still left unanswered. First, how important is a self-selection effect compared to an incentive effect? To separately quantify both effects requires us

to develop a new experimental design in order to create exogenous variations in a self-selection process. Second, can we regard the correlation between social preferences and individual sorting behavior as the causality? Identifying the causal relationship between social preferences and individual behavior is notoriously difficult due to endogeneity problems. Third, in order to ensure the external validity, we should accumulate additional experimental evidence for larger samples, more realistic workplaces, and more relevant social preferences such as fairness or sensitivity to social reputation than the controlled lab experiment by Dohmen and Falk (2011). Finally, if social preferences and norms are effective in workplaces, under what conditions do they emerge? Repeated interaction and informal enforcement mechanisms such as reputation seem to play significant roles in laboratory experiments (Bó, 2005; Bó & Fréchette, 2011), but, to the best of our knowledge, few empirical studies provide the answer for this question due to the difficulties of collecting relevant data.

This paper is the first attempt to give answers to all of above questions. To do so, we developed and implemented the two-stage field experiment in the traditional rice-planting contract in the Philippines. We recruited a pool of rice-planting workers and organized them into several working teams, each of which consists of four to seven workers. In the first stage of the experiment, we offered two options of contracts for workers, i.e. fixed wage (FW) and individual piece rate (IPR), and each worker could choose one of them that he or she prefers. In the second stage, workers who chose IPR were randomly converted into FW with 50% probability. Then they were asked to plant rice seedlings in assigned plots together with own team members. These processes constitute one experimental session and our data includes 28 sessions. In addition, we carried out two types of lab-in-the-field experiments for all of participants in the field experiment, i.e. ultimatum game and kinship tax game to elicit inequity aversion (Fehr & Schmidt, 1999) and kinship tax rate (Squires, 2016), respectively.[1] These experiments are modified versions of standard ultimatum game and kinship tax game adopted by Beekman, Gatto, and Nillesen (2015); Boltz, Marazyan, and Villar (2015); Squires (2016) such that each subject is matched with *all* of subjects except for oneself. Therefore elicited individual preferences and kinship tax rate are varied with the matched partners.

Each stage in the experiment is well-designed to answer specific research questions. First, comparing productivity of those who chose FW in the first stage to those who

[1]Kinship tax rate can be defined as the willingness-to-pay to hide own income from specified others. Hence, kinship tax rate reflects the strength of social pressure to redistribute one's own income to socially connected people.

chose IPR in the first stage but were converted into FW in the second stage allows us to accurately measure a self-selection effect. An incentive effect can also be quantified by comparing productivity of those who chose IPR but were converted to those who chose IPR but were not converted. Second, the random assignment of peers in every session, combined with the data of social preferences and kinship tax rate for each peer, permits us to identify the causal impacts of social norms on workers' choice decisions.[2] Third, the randomized team formation procedure also makes it possible to utilize the fact that seasonal workers temporally migrate to the survey villages. Specifically, we randomly assign the seasonal workers into teams which are composed by local workers. Since they are not socially connected each other at all and seasonal workers never return back to the same village at least for several years this creates exogenous variations in terms of the intensity of repeated interaction within teams. Then we examine the causal impacts of the (possibility of) infinite repeated interaction on the emergence of social norms in worker's self-selection process. This can be considered as the first experimental investigation for the community enforcement mechanism suggested by Kandori (1992).

There are three main empirical results. First, a simple econometric analysis shows that the self-selection effect is more prominent by 1.6 times than the incentive effect in explaining the productivity gap between FW and IPR. Second, social norms shape a major part of the self-selection process. Guilt aversion and kinship tax can be obstacle for workers to opt for more remunerative option while enviousness pushes them to choose IPR to reduce an income disparity between teammates. Third, social norms are likely to emerge if workers conceive that the infinite repeated interaction and the informal reputation system are in effective.

This paper contributes several strands of the literature. First, this paper is closely related to the literature that separately identifies the selection and incentive effects in the workplaces. As the seminal work, Lazear (2000) quantifies worker-selection effects using non-experimental panel data collected in the large manufacturing factory in the U.S., as opposed to causal effects of financial incentives, and concludes that a half of the increase of productivity might result from the sorting into the piece-rates scheme by more abled workers. Some recent papers utilizing experimental methods confirm the Lazear (2000)'s conclusion on the importance of selection effects (Guiteras & Jack, 2014; Leuven, Oosterbeek, Sonnemans, & Van Der Klaauw, 2011; Cadsby, Song, & Tapon, 2007). Among other things, the most relevant paper is Kim, Kim, and Kim (2017). They

[2]Here, we define social norms as the sum of bilateral social preferences toward teammates. The detailed definition is elaborated below.

provide empirical evidence on how career and wage incentives affect labor productivity through worker selection and causal effect channels by the two-stage randomized controlled trial in Malawi. In their setting, the career incentives are the combination of a future job prospect and a recommendation letter while the wage incentives are a lump-sum salary and a performance-based bonus payment. Our paper extends Kim et al. (2017) in three ways. First, we focus on more general and broadly-common incentive contracts which are the same schemes with Lazear (2000). Second, we implemented multiple experimental rounds in each of which workers were repeatedly asked to select out one of contracts from available options, which enables us to estimate how persistent the effect of the choice of contracts is and how much their behavior is path dependent. Third, we explore the role of social norms as key explanatory variables to determine the selection process on which Kim et al. (2017) shed less light.

Second, this paper compliments the another literature on the role of social norms in shaping economic behavior (see Festre (2010) for a comprehensive review). Specifically, our paper precisely identifies the causal impacts of social norms on the self-selection process in the agrarian labor contract. Third, this paper extends theoretical literature on cooperative behavior in the repeated game and informal enforcement mechanisms by investigating under what conditions social norms emerge in village communities. Exogenous variations of the repeated interaction and the availability of past history of matched players quantify the causality. Finally, this paper also contributes the methodology of field experiments by showing that a two-stage field experiment can parse out the selection process from incentive effects such as Ashraf, Berry, and Shapiro (2010), Karlan and Zinman (2009), and Kim et al. (2017), employing multiple stages of randomization.

The rest of paper is structured as follows. Section 2 describes the context and experimental designs. Section 3 and 4 explain the data and show descriptive statistics for key variables. Section 5 provides econometric strategies and empirical results. Section 6 concludes.

2 Design of Experiments

We analyze behavior of rice planting workers in two rural villages in the Central Luzon in the Philippines during the 2016 dry season. There are four data sources. The first one is the two-stage field experiment to identify the incentive and self-selection effects. The second one is the lab-in-the-field experiments to elicit worker's social norms. The third one is the household questionnaire survey to capture the economic and social characteristics

of workers. The fourth one is the social network survey in which the complete social networks within the communities are depicted.

2.1 Context

In a season of rice planting, the arrangement and management of daily workers for planting seedlings are indispensable since it will affect the timing to fill water in a paddy field before land protection, which crucially determines a yield level in a coming harvest season. However, it is virtually impossible that a farmer by himself organizes a plenty of workers to complete a rice planting task properly. Therefore, there are loosely-tightened organizations which are managed by local supervisors. The structure of these organizations and labor contracts with farmers are as follows. First, a farmer requests a supervisor to arrange a pool of workers (approximately twenty workers per hectare) and pays a certain amount of cash which is proportional to an area of a paddy field. Second, a local supervisor takes the responsibility to complete rice planting and sends workers to the contracted field in a given day. Third, after finishing a task, a local supervisor shares a payment with workers who participated in it. So the amount of payment per person is usually fixed at 100PHP to 120PHP and common among all workers. Now we discuss below some key features to understand the main empirical contexts.

Key Feature 1: Team Production. In the field site, a pool of workers is assigned into sub-groups each of which consists of four to six workers. Each sub-group is responsible for one to two paddy fields. The members of each sub-group are supposed to plant together and not allowed to work with other members who belong to different sub-groups. The workers' primary task is merely to plant seedlings. The only choice variable of workers is how much effort to exert into planting in this context. Each worker plants seedlings individually in distinct rows and the default payment scheme is the fixed wage contract. However, a principal or a supervisor sets the minimum acceptable amount of group output for each team, below which a supervisor rejects to pay for workers. Therefore, one's effort decreases the amount of tasks that other teammates are required to exert. So there exist the externality between workers arising from the production technology as well as the payment scheme. There are two sessions in a day, that is, the morning and the afternoon. In every session sub-groups are reorganized.

Key feature 2: The Characteristics of Task. The task of rice planting is remarkably simple, that is, merely to plant seedlings in a straight way. This characteristics has three

advantages. First, the outcome or productivity which can be defined as the length of planted lines in a unit of meter per 10 min is easily and accurately observed by other teammates, local supervisors, and experimenters. Second, the rice planting is not the multi-task problem (Holmstrom & Milgrom, 1991) which includes the trade-off between quantity and quality. Umekage, Inoue, and Watabe (1966) demonstrates the rice planting is too simple to manipulate the quality of task.[3] Third, the simpleness implies that there are no learning effects even between the beginner and the veteran, which means that the different productivity reflects the difference in individual effort levels.

Key feature 3: The Coexistence of Local and Seasonal Workers. A local supervisor has generally the responsibility to manage two types of pools of workers. Local workers live in the same village with a local supervisor and the others come to and temporally stay the village only during a planting season. Seasonal workers migrate from the various villages and usually they never come back to the same village. This is because they can get permission to migrate to the given village only for a single season and which villages they will migrate to in a next season depends on the referral prepared by the current local supervisor. The current local supervisor introduces them to the another supervisor who works for and manages a different pool of worker's organization in a different village. Usually, local workers have no interaction with seasonal workers both in the daily life and in the workplace. This situation provides the exogenous variation in terms of the possibility of repeated interaction among teammates by mixing those workers randomly. We use this variation to prove empirically whether infinite repeated interaction is the precondition of the emergence of social norms in the workplace.

2.2 Field Experiments

We implement the two-stage field experiment which is aimed at identifying both incentive and self-selection effects at the same time in the workplace. We explain the procedures and the structure of this field experiment. Fist, the experimenter and the local supervisor

[3]Umekage et al. (1966) carries out experiments to identify the effect of various transplanting methods on the growth habits of rice plants by comparing two methods: planting rice seedlings in a horizontal direction and in a vertical direction (which is the normal method). They find that the difference in the yields between the plots is not significant at the 5% level. Moreover, according to the local supervisor, not only the speed but also the quality of planted rice can be easily observed and verified by checking the misalignment and lodging of seedlings. Hence, we believe that multi-task problems are not necessarily serious in our experiments.

recruit workers who can participate in an experimental session. Second, the experimenter gives the instruction and confirms their agreement to the contents of the experiment. Third, the experimenter randomly forms sub-groups or teams which are each composed by four to six workers. Note that this randomized team formation doesn't distinguish local and seasonal workers and then we have exogenous variations in terms of the proportion of local (or seasonal) workers within a team. During an experimental session which takes 30 min workers are required to plant seedlings in assigned plots with teammates and are prohibited to work with the other workers who belong to different teams.

Figure 1 shows the structure of the two-stage field experiment. In the first stage, the experimenter proposes two types of contracts (i.e. FW or IPR) to the workers. Contracts are as follows:

$$\text{FW}: \ w_{ij}^{FW} = \begin{cases} F & \text{if } \sum_{k \in N_j} l_{kj} \geq \underline{l} \\ 0 & \text{if } \sum_{k \in N_j} l_{kj} < \underline{l} \end{cases} \tag{1}$$

$$\text{IPR}: \ w_{ij}^{IPR} = \begin{cases} F + \alpha(l_{ij}) & \text{if } \sum_{k \in N_j} l_{kj} \geq \underline{l} \\ 0 & \text{if } \sum_{k \in N_j} l_{kj} < \underline{l} \end{cases} \tag{2}$$

were l_{ij} is the observed outcome of worker i in group j. $\underline{l}$ is the minimum requirement level of total outcome for each group below which every workers in group j cannot get paid. Then the workers choose one of the contracts under which they prefer to work. We set that F equals to 40 PHP and $\alpha = 5$.

In the second stage, the experimenter randomly converts some workers who opt for IPR to FW with 50% probability. There might exist the disappointed effect: the workers put less effort in the converted FW contract compared to the case that workers are not allowed to choose any contract by themselves and are randomly assigned into FW. To minimize this disappointed effect, the experimenter explains the procedures and the possibility of forceful conversion clearly. After conversion, workers are asked to engage in the rice planting task and the enumerators measure individual productivity which is defined as the length of planted lines per 10 min.[4] After an experimental session, the workers get paid immediately based on their contract and productivity. The information on the individual payment and the contract choice is shared among teammates but strictly

[4]The enumerators trace all movement by each worker inside the plot by drawing the traffic lines on the working sheet. In order to assure the precision of measurement of productivity we conducted pilot experimental sessions to train the enumerators. Additionally, we used a video camera to capture all of traffic lines of each worker and confirms that there is no difference in individual productivity among the different methods.

confidential to the other workers.

2.3 Lab-in-the-field Experiments

Three types of lab-in-the-field experiments were carried out in July 2016. Those are ultimatum game, kinship tax game, and risk game. Subjects were recruited from the pool of participants in the field experiment. We held the instruction session for all subjects at the school cafeteria, which took 4 hours in total. After this session, subjects got back home and trained enumerators visited them to conduct experimental sessions individually. Three experiments took approximately five hours in total for each subject. Detailed design for each experiment is described below.[5]

2.3.1 Ultimatum Game

We conduct a modified version of ultimatum game which is aimed at structurally estimating proposers-and-responders-specific preferences for inequity aversion. It is worth noting that although previous experimental studies usually provide a single parameter for inequity aversion that are common for all subjects our experimental design allows us to structurally estimate parameters differing among not only proposers but also their responders matched in the game.

In this game, a proposer is endowed with 100PHP and asked to decide how much to share with a responder. The proposer's choice set is discretized to eight allocations:

$$A \in \{(100, 0), (85, 15), (70, 30), (55, 45), (45, 55), (30, 70), (15, 85), (0, 100)\},$$

where the first and second amounts denote the offer for a proposer and a responder, respectively.

We purposefully rule out the equal sharing case to force proposers to commit themselves to offering either more or less than the equal split, which should increase the efficiency of our structural estimates. A responder could decide whether to accept or reject a proposer's offer. Here, the strategy method is adopted, which means that a responder is asked to answer for each of the eight allocations. Note that in our design all subjects are required to play as both roles. In addition, a proposer should answer on her offers for each of possible responders which consists of all workers who are invited to the experiments. In the same manner, a subject is also ordered to answer on her decisions as a responder for

[5]In terms of risk game, we adopted the identical design with Holt and Laury (2002).

each of possible proposers. In other words, each subject is matched as both of a proposer and a responder with all subjects.

As Bellemare, Kröger, and Van Soest (2008) discussed, in order to precisely estimate parameters for inequity aversion, we should collect the data on proposers' subjective probability distributions over the actions of responders rather than relying on the assumption that a proposer has rational expectations. The beliefs of proposers are elicited with a series of subjective probability questions. Subjects are asked what the probability that each responder accepts each offer would be.[6]

Decision-making is incentivized by real money: after all decision is made by subjects an experimenter randomly form the pairs and relevant transfers are exercised, which is clearly mentioned in the instruction session.

2.3.2 Kinship Tax Game

In order to directly measure kinship tax rates for each subject, we adjust the existing laboratory experiments which are used in Beekman et al. (2015), Boltz et al. (2015), and Squires (2016).

In our experiment the Becker-DeGroot-Marschak (BDM) mechanism is adopted to elicit willingness-to-pay (WTP) to hide income from a social network or to escape social pressure that obligates individuals to commit informal income redistribution. The rate of this WTP to a total endowment is defined as the individual kinship tax rate. The BDM method that operates much like a second-price auction against an unknown or random price might be more suitable in our particular context than the method of take-it-or-leave-it (TIOLI) offers in which single randomized prices are assigned to each individual. This is because although TIOLI has a simple structure to understand and implement it provides limited information which only includes a subject's dichotomous decision-making on whether or not a subject buys a certain good at the assigned price, while the BDM mechanism elicits subject's exact WTP for a good. This precise measure of WTP is necessary to quantify the impacts of kinship tax on the decision-making regarding the contract choices.

In our design, there are three sizes of initial endowment, i.e. 40PHP, 120PHP, and 360PHP, and two distinct framing as to where the endowment come from, i.e. earned and windfall income. Therefore all subjects are asked to play six different kinship tax games.

[6]To be able to account for framing effects, proposers are randomly divided into groups that are asked for either their subjective acceptance or rejection probabilities for all offers.

Since the structure of the game is identical between three sizes of initial endowment we will give detailed explanations for the two types of framing in terms of income sources.

The Earned-income-framing. There are three procedures in this framing. First, the experimenter gives a detailed instruction in a common room. All subjects are informed that they would have a chance to earn a reward which is equal to for example 120PHP by completing the task within the experiment. The task is simple such that all subject could easily understand, which is to sort the cards in order of increasing number within certain minutes.[7] If she could successfully complete the task then she would get a reward, but an amount of a reward was supposed to be announced in public.[8] The experimenter demonstrates in advance how to play the card sorting game and how to reveal an amount of rewards.

Second, after an instruction, the enumerators call each subject to the decision rooms that are isolated each other such that a subject can not observe other subjects' decision and vice versa. she is asked to answer her maximum WTP to hide the income from other participants: how much cost at maximum could she afford to make her income confidential?[9]

Third, after the elicitation of WTP, she engages in the task. If she successfully completes it the price is randomly drawn by the experimenter. If the random price is lower than or equal to her WTP she is able to buy the right to hide the income at that random price otherwise she would pay nothing and her income would be disclosed to the partner. If she fails to accomplish the task she is not permitted to get any reward.

The Unearned-income-framing. The instruction is identical to the last framing except that income would be coming from the lottery in this framing. Instead of doing the task the lottery is drawn by a subject after the elicitation of WTP. The probability to win the prize is fifty percent. If she could win the prize the price to hide income is randomly picked up by the experimenter. Then the same procedures with the last framing are applied. In order to avoid learning and ordering effects the sequence of two framing

[7] In the card sorting task, a subject could get paid if she is able to sort twenty cards in order of number within a reference time. A reference time was stratified by age, educational attainment, and sex of each subject so as to set the probability to complete the task at 50 percent within each cohort. The contents and the reference time are clearly mentioned before their decision with regard to WTP.

[8] Precisely, her reward would be disclosed to each matched participant in a given session in a way that the partner could identify both her name and an amount of a reward.

[9] WTP could be answered by a unit of 1 Centavo which is equal to 0.01PHP.

is randomized within a session.

This kinship tax experiment has several features that previous studies does not explicitly deal with. First, unlike other kinship tax games, our experiment is designed to draw out payers-and-receivers-specific kinship tax rates. This means that potential tax payers are asked to answer their WTP to hide income from each matched partner. This enables us to estimate WTP varied with the $N_j - 1$ matched partners where N_j is the number of participants in a session j.

Second, Boltz et al. (2015) and Squires (2016) that adopt some sort of BDM mechanism might not be able to eliminate biases in estimating WTP due to the following reasons: (a) subject's behavior can possibly be distorted by the sets of choices in their designs since they show that the peaks in the distribution of individual WTP are at the endpoints of possible choices regardless of the actual tax rates at those points:[10] (b) the multiple choice list design which is the most common method in existing literature might underestimate true WTP since some subjects should fall in a range of choices and in this case WTP is calculated as the lower bound. To minimize these biases we ask the exact amount of WTP for each subject which is not discretized by the certain choice set and we do not explicitly set the endpoint. In addition to this, we also conduct the same experiment with Squires (2016) to replicate his results and compare with the ones being observed in our design.

Third, the most critical issue in the literature is the problem of the pseudo-zero tax rate. In a design where WTP to deviate secretly from the income redistribution norms is defined as kinship tax rate, the subject who considers paying the tax to the others as duty would answer zero WTP since she feels obligated to commit this arrangement.[11] Therefore even though the WTP should strictly be positive for this subject the BDM mechanism could not appropriately capture this phenomenon. In order to disentangle the true tax rate from this disguised zero tax rate the subject who answers zero WTP at the first decision stage is asked to answer the following question: what is the tax ratio

[10]Boltz et al. (2015) demonstrates that conditional on positive WTP more than fifty percent of their sample chooses the corner solution which is equal to 12.5% kinship tax rate. Squires (2016) specifies that more than fifty-three percent of his sample is on the edge of the choice set where WTP to hide income is 70%. Bohm, Lindén, and Sonnegård (1997) find that bids in the BDM are sensitive to the choice of endpoints of the distribution of possible transaction prices.

[11]Note that the dominant strategy for the subject who thinks that the tax payment was duty is to declare zero WTP. This is because that even if she hides her reward by paying some sort of money she is obligated to pay the tax which would be proportional to the initial endowment i.e. 40PHP, 120PHP, or 360PHP. Therefore there is no incentive to hide her income from the others.

that your matched partner would think that you have to send to her? Then, we ask the matched partner "What do you think is the tax rate that your partner should share with you?". If both answers are close enough the additional reward is paid for both players. In the case that both answers show positive kinship tax rate and are almost identical each other we replaces zero kinship tax with this positive value.

3 Data

We conducted the field experiments for 141 workers from two villages. The 73 workers came from M-village and 68 workers came from G-village and all of them participated in lab-in-the-field experiments, which are including 25 and 24 seasonal workers, respectively. The total number of rounds for field experiments are 16 in M-village and 12 in G-village. Total observations are 2,352.

Table 1 shows the individual productivity defined as the planted length per 10 min by contracts. The workers achieve the highest productivity under IPR which is 22.55 m/10min followed by the workers who are randomly converted from IPR under which the productivity is 19.41 m/10min. The least productivity is recorded in FW that is 14.74 m/10min.

Figure 2 shows the cumulative distribution functions (CDFs) by contract schemes. This figure demonstrates two facts. The first one is that there exists the threshold in terms of the minimum requirement effort level. Especially, the shape of CDF for FW implies that the distribution of individual productivity in this contract is positively skewed and cut off nearly at the minimum requirement effort level. The second one is that both incentive and self-selection effects are graphically observed. we can verify that the CDF of the IPR contract dominates that of the FW and the converted FW contracts. The two-sample Kolmogorov–Smirnov tests of the equality of distributions reject equality between these three CDFs at the 1% level of statistical significance.

4 Construction of Key Variables

4.1 Structural Estimation for Inequity Aversion

In this section, we introduce a structural econometric model to estimate the parameters for inequity aversion which are specific for each matched pair. The model that we exploit is based on Bellemare et al. (2008) although it is followed by some modifications.

4.1.1 Preferences Varied with Dyadic Relationships

Suppose that subjects have proposers-and-responders-specific preferences with possibly nonlinear asymmetric inequity aversion. The utility of proposer i who is matched with responder k is given by

$$\begin{aligned} u_{ik}(x) = x_{ik} - \alpha_{1ik} \max\{x_{ki} - x_{ik}, 0\} - \alpha_{2ik} \max\{x_{ki} - x_{ik}, 0\}^2 \\ -\beta_{1ik} \max\{x_{ik} - x_{ki}, 0\} - \beta_{2ik} \max\{x_{ik} - x_{ki}, 0\}^2, \end{aligned} \tag{3}$$

where x_{ik} is monetary payoffs for i who is matched with k and vice versa.

The linear inequity aversion of Fehr and Schmidt (1999) is a special case of equation (3) with $\alpha_{2ik} = \beta_{2ik} = 0$.

I use following specifications:

$$\alpha_{1ik} = \exp(z'_{ik}\alpha_1 + \nu^{\alpha}_{ik}), \tag{4}$$

$$\beta_{1ik} = \exp(z'_{ik}\beta_1 + \nu^{\beta}_{ik}), \tag{5}$$

$$\alpha_{2ik} = z'_{ik}\alpha_2, \tag{6}$$

$$\beta_{2ik} = z'_{ik}\beta_2, \tag{7}$$

where z_{ik} is a vector of the differences and the sums of observed characteristics between i and k, and the attributes between them such as social and geographical proximity. Since the preference parameters have directed dyadic structures observed characteristics that are supposed to affect preferences' distribution must be entered in a way that the symmetric condition is satisfied, which requires that the variables (z_i, z_k) should affect the parameters (e.g. α_{ik}) in the same way that the variables (z_k, z_i) affect the opposite dyadic parameters (e.g. α_{ki}). ν^{α}_{ik} and ν^{β}_{ik} reflect unobserved heterogeneity that is varied with not only the proposers but also the responders. We assume that both ν^{α}_{ik} and ν^{β}_{ik} are independent of error terms and of z_{ik} with a bivariate normal distribution with means zero and an arbitrary covariance matrix.

Decisions of Proposers in the Ultimatum Game. Each proposer had eight choices $(j = 1, \ldots, 8)$, involving own payoffs $x_{ik}(1), \ldots, x_{ik}(8)$. Subjects are assumed to maximize their expected utility, where proposer i uses the own subjective probability Q_{ikj} that offer j will be accepted by k. Since utility is zero if the offer is rejected, the expected utility of offer j is given by $Q_{ikj}u_{ikj}$, where u_{ikj} denotes person i's utility of payoffs $(x_{ik}(j), 100 - x_{ik}(j))$.

The subjective expected utility of making an offer $x_{ik}(j)$ is therefore given by

$$\begin{aligned} Q_{ikj}u_{ikj} = Q_{ikj}[x_{ik}(j) &- \alpha_{1ik}\max\{100-2x_{ik}(j),0\} \\ &-\alpha_{2ik}\max\{100-2x_{ik}(j),0\}^2 \\ &-\beta_{1ik}\max\{2x_{ik}(j)-100,0\} \\ &-\beta_{2ik}\max\{2x_{ik}(j)-100,0\}^2]. \end{aligned} \tag{8}$$

To allow for suboptimal behavior, we add idiosyncratic error terms ϵ_{ikj} multiplied with a noise-to-signal ratio parameter λ_{ik}. we assume that errors ϵ_{ikj} are independent of each other and of other variables in the model (i.e. $(\nu^{\alpha}_{ik}, \nu^{\beta}_{ik})$, and z_{ik}), and that the difference of any two ϵ_{ikj} across options follows a logistic distribution. We assume that proposer i chooses the option j that maximizes $Q_{ikj}u_{ikj} + \lambda_{ik}\epsilon_{ikj}$.

Decisions of Responders in the Ultimatum Game. Responder i is supposed to decide to accept or reject each offer based on the her utility. The utility of rejecting is zero, and the responder utility u_{ikj} of accepting offer j immediately follows from equation (18):

$$\begin{aligned} u_{ikj} = x_i(j) &- \alpha_{1ik}\max\{100-2x_{ik}(j),0\} \\ &-\alpha_{2ik}\max\{100-2x_{ik}(j),0\}^2 \\ &-\beta_{1ik}\max\{2x_{ik}(j)-100,0\} \\ &-\beta_{2ik}\max\{2x_{ik}(j)-100,0\}^2. \end{aligned} \tag{9}$$

We assume the responder accepts offer j if $u_{ikj}+\lambda_{ik}\epsilon_{ikj} > 0$. We allow the noise parameter λ_i to vary with observed characteristics by assuming that $\lambda_{ik} = \exp(z'_{ik}\lambda)$.

4.1.2 Beliefs of Proposers

The observed subjective probabilities are assumed to be generated by the following process:

$$P^*_{ikj} = z'_{ik}\delta + \gamma_j + \nu^P_{ik} + \epsilon^P_{ikj}, \tag{10}$$

$$P_{ikj} = \begin{cases} 0, & \text{if } P^*_{ikj} \leq 0, \\ P^*_{ikj}, & \text{if } 0 < P^*_{ikj} < 1 \\ 1, & \text{if } P^*_{ikj} \geq 1. \end{cases} \tag{11}$$

Since true as well as reported probabilities may well be 0 or 1, we allow for censoring at 0 and 1, as in a two-limit tobit model. The choice option effects γ_j are expected to

increase with j for amounts bellow the equal split, since proposers probably realize that acceptance probabilities rise if the amount offered to the other player increases toward an equal split.

The correct process generating proposer expectation is assumed to be

$$Q^*_{ikj} = z'_{ik}\delta + \gamma_j + \nu^P_{ik}, \tag{12}$$

$$Q_{ikj} = \begin{cases} 0, & \text{if } Q^*_{ikj} \leq 0, \\ Q^*_{ikj}, & \text{if } 0 < Q^*_{ikj} < 1 \\ 1, & \text{if } Q^*_{ikj} \geq 1. \end{cases} \tag{13}$$

We assume that the triplet $(\nu^\alpha_{ik}, \nu^\beta_{ik}, \nu^P_{ik})$ is distributed as a trivariate normal distribution with arbitrary covariance matrix, independent of observed characteristics and other error terms in the model.

4.1.3 Parameter Estimates

We estimated the model in section 5.1.1 and 5.1.2 by maximum simulated likelihood.[12] Table 2 shows the estimation results of the model. There are four findings that should explicitly be noted. First, the average estimated parameters for individuals are 0.281 for α_i and 0.135 for β_i which are plausibly in line with the theoretical assumption in Fehr and Schmidt (1999). It implies that our estimated model would be suitable. Second, in both parameters α_1 and β_1, several observed characteristics are statistically significant; the difference between and the sum of variables for matched subjects regarding educational attainments, age, and income. For example, the difference of income variable has a negative coefficient in α_1 which is statistically significant at 5% level. This means that a subject will have disutility from her disadvantage when the income level is close with her matched partner. These estimation results imply that the preference for inequity aversion differs among not only the proposers but also her matched responders.

Third, interestingly, inequity aversion for both one's advantage and disadvantage is likely to be varied with social and geographical connections. For instance, people tend to dislike others' superiority if the others are close enough to them. The parameters α_2 and β_2 are also affected by the social and geographical proximity: the setting of nonlinear preferences which are correlated with the intensity of social interactions are more appropriate in this setting than linear ones.

[12] Detailed estimation procedures and the anatomy of the determinants of inequity aversion are provided in the appendix.

In the main empirical analysis, we will utilize these proposers-and-responders-specific preferences on inequity aversion regarding both one's advantage and disadvantage as key explanatory variables to quantify the causal impacts of social norms.

4.2 Estimates of Kinship Tax Rate

Table 3 and Figure 3 show the estimated social reputation tax rates and their distributions which are based on the experimental results. There are several findings which are worth noting. First, the average tax rate is 2.38% for the whole observations, 3.67% for the observations conditional on positive tax rate. Furthermore, the tax rate would plausibly fall into the range between 0 to 14.5% without extreme outliers. These percentages seem to be moderate, compared to the other estimates reported in existing literature. This fact implies that our experiment would precisely be able to measure tax rates.[13]

Second, more than 50 percent of all observations indicate positive social reputation tax rate. Specifically, the percentages for small, medium, and large initial endowment in the case of earned income are 53.2%, 52.8%, and 58.8%, respectively and 60.0%, 59.6%, and 63.5% in the case of windfall income, respectively. In addition, a subject expresses positive tax rates for approximately a half of a pool of participants in a given session. Therefore, it is evident that kinship tax rate is prevalent among subjects so that it would affect their decision-making or behavior in the labor contract that we investigate below. Moreover, obtaining windfall income would more prompt people to engage in the arrangement of income redistribution than earned income.

Third, the average tax rate increases with the size of initial endowment. In all sample, the average tax rates for small, medium, and large endowment are 1.94%, 2.02%, and 2.29%, respectively. Conditional on the positive tax rate, this trend can also be confirmed: 3.42% for small, 3.54% for medium, and 4.05% for large endowment, respectively. To our knowledge, this is the first evidence that kinship tax rates will positively be correlated with the size of initial endowments. This fact indicates that kinship tax might be sensitive

[13]In Appendix we can accurately replicate the results presented by Squires (2016) when we use the data coming from his design but targeting on the same sample. This replication and an additional analysis imply that the preference for ambiguity aversion explains the extremely high tax rate (53.1 percent of the sample are categorized into the group who has 70% tax rate conditional on positive tax rate) in his design. When we counterfactually set the impact of ambiguity aversion on the tax rate at zero for all sample the predicted tax rates from econometric estimation are almost identical within the same individuals regardless of the difference of experimental designs. For more detailed discussion please see the Appendix.

to the wealth or relative economic status between the tax payer and the receiver.

Fourth, the average tax rates for windfall income are always higher than ones for earned income regardless of amounts of initial endowment. This implies that whether or not a subject deserves to obtain the reward is a critical determinant of the level of tax rates. Interestingly, when we look at the distribution of kinship tax rate by the initial endowment sizes, the two-sample Kolomogorov-Smirnov tests reject the equality of the cumulative distribution functions (CDFs) between earned and windfall income at the 1% level of statistical significance in the medium endowment while the equality hypothesis of the CDFs between them in the small and large endowment cannot be rejected even at the 10% level. This statistical result strongly suggests that if the stake is relatively small or large enough people are less likely to take account of the deservedness of the reward. In contrast, this deservedness are seriously taken into account in the medium endowment which is equal to the level of daily income and tax allowance is introduced for earned income. Based on these results, we will use the kinship tax rate for the earned income in the medium endowment as the key explanatory variable for the main empirical analysis regarding the decision-making in the labor contract.[14]

5 Empirical Analysis

5.1 Self-selection Effect vs. Incentive Effect

5.1.1 Identification

The design of the two-stage field experiments allows for parsing out the self-selection effect from the incentive effect. In order to compare the scales of these two effects and decompose them by observed covariates, we consider the following equation:

$$y_{ijt} = \alpha^{base} + \beta^{base} IPR_{ijt}^{choice} + \gamma^{base} IPR_{ijt}^{contract} + \epsilon_{ijt}^{base} \tag{14}$$

where y_{ijt} is worker i's productivity which is defined as the length (in a unit of meter) of planted line per 10 minutes in group j at round t. IPR_{ijt}^{choice} is the dummy variable taking one if a worker chose IPR contract in the first stage of the decision-making and zero otherwise. $IPR_{ijt}^{contract}$ is the dummy variable taking one if a worker was actually engaged in rice planting under IPR contract and zero otherwise. The parameters of interest are β and γ which capture self-selection and incentive effects, respectively. More formally, Δ^S

[14] Also, robustness checks with different definitions are conducted in the Appendix.

denotes the self-selection effect:

$$\begin{aligned}\Delta^S &= E[y|IPR^{choice}=1, IPR^{contract}=0] - E[y|IPR^{choice}=0, IPR^{contract}=0] \\ &= \alpha^{base} + \beta^{base} - \alpha^{base} = \beta^{base}. \end{aligned} \tag{15}$$

In the same way, we can derive the incentive effect or Δ^I:

$$\begin{aligned}\Delta^I &= E[y|IPR^{choice}=1, IPR^{contract}=1] - E[y|IPR^{choice}=1, IPR^{contract}=0] \\ &= \alpha^{base} + \beta^{base} + \gamma^{base} - (\alpha^{base} + \beta^{base}) = \gamma^{base}. \end{aligned} \tag{16}$$

In other words, Δ^S measures the mean difference of productivity between the workers who chose IPR and FW conditioned that both of them finally got paid under FW scheme whereas Δ^I measures the mean difference of productivity between the workers who were engaged in IPR and FW conditioned that both of them chose IPR in the first stage.

Note that this specification allows β to be affected by both observable and unobservable covariates which are included in the residual term. This parameter is exactly what our experimental design aims at estimating. Put differently, it is one of the main purposes of the paper to show the productivity gap caused by self-selection by permitting both observable and unobservable covariates to contribute this gap. In the following section, we furthermore decompose this self-selection effect by observable covariates. Consequently, $|\Delta^S|/|\Delta^I|$ indicates the relative importance of the *aggregate* self-selection effect on worker's productivity compared to the incentive effect which can precisely be identified because $IPR_{ijt}^{contract}$ is perfectly exogenous conditional on IPR_{ijt}^{choice}.

5.1.2 Baseline Results

Table 4 presents the estimation results of the above specification. The fist column shows that the self-selection effect or the mean productivity gap between the workers who chose IPR in the first stage but were converted into FW and who chose FW in the first stage is 4.67 meter per 10 min and statistically significant. This is equal to 31.7% of the average productivity in FW. On the other hand, the coefficient of a dummy variable or $IPR_{ijt}^{contract}$ is 3.11 and statistically significant, which implies that the monetary incentive increases a worker's productivity by 16.2%. The difference between these coefficients is statistically significant ($p < 0.01$) and $|\Delta^S|/|\Delta^I| = 1.502$. Therefore, although the extrinsic incentive significantly affects worker's behavior through both effects, the self-selection effect is more prominent in our setting than the incentive one. Compared to Lazear (2000)'s estimation,

the scales of both effects seem to be relatively low but the ratio of self-selection effect to incentive effect is reasonably same with his result.[15]

In the second column, we focus on the choices which were made by workers who participated in the experiment for the first time. Still we observe significant self-selection and incentive effects for this sub-sample, which implies that even for workers who are less likely to know others' belief and behavior self-selection effects are robustly prominent.

5.1.3 De-motivated by Random Conversion?

One might be concerned that our experimental design which forcefully converts a worker who choose IPR to FW might disappoint a converted worker. This *de-motivation effect* can create biases in above estimations for self-selection effects. Note that the de-motivation effect decreases the self-selection effect and thus our estimated effect can be considered as the lower bound. Table 5 documents the estimation results to check whether the de-motivation effect is problematic in our context. The first column replicates the last equation's result. In the second column, we compare those who are forcefully converted into FW for the first time to those who are converted more than once. The de-motivation hypothesis predicts that the former feels more disappointed since he does not get used to the random conversion. So, the interaction term between the first conversion dummy and the self-selection effect should have a negative coefficient in this case. Our estimation result rejects this hypothesis since the coefficient is not statistically significant. Next, we also conduct additional field experiments where the conversion rate from IPR to FW is 0.3 rather than the default rate or 0.5. The lower conversion rate will exacerbate to what extent he gets disappointed by forceful conversions. However, according to the third column, this is not the case in our context. The fourth column demonstrates that the de-motivation effects do not exist even if a given worker is the only person who is converted to FW within a team. Finally, the worker who can earn more money in IPR than others do is subject to suffer the larger negative impact of the forceful conversion. Even for those workers, we can not confirm de-motivation effects. In sum, we find no evidence on the existence of the de-motivation effect in our context.

[15]Lazear (2000) found that the total productivity gain by switching from FW to IPR is 44% of which a sorting of workforce would be explaining approximately 22%.

5.2 Who Opts for Individual Piece Rate?

In this section, we present the evidences to answer what sort of factors shape the self-selection process and how important social norms are in this process.

5.2.1 Econometric Specification

In order to rigorously identify the impacts of covariates on a worker's choice, there should be exogenous variations in terms of them. The random assignment of co-workers in forming a team combined with the parameters of social preferences varied with the matched partners allows us to precisely quantify the impacts of social norms on an individual choice. Consider a following regression model:

$$\begin{aligned} IPR_{ijt}^{choice} = \kappa + \phi \textstyle\sum_{k\in n_j, k\neq i} Kinshiptax_{ijt}^{k} + \mu \sum_{k\in n_j, k\neq i} Envy_{ijt}^{k} \\ +\gamma \textstyle\sum_{k\in n_j, k\neq i} Guilt_{ijt}^{k} + \tau \bar{\hat{a}}_{-ijt} + X'_{ijt}\nu + \psi RepInteract_{ijt} \\ +\omega PastHistory_{ijt} + \rho_i + \eta_{jt} + \xi_{ijt}, \end{aligned} \tag{17}$$

where $Kinshiptax_{ijt}^{k}$ is a kinship tax rate that worker i is obligated to pay co-worker k. n_j is a total number of workers in group j. Therefore, $\sum_{k\in n_j, k\neq i} Kinshiptax_{ijt}^{k}$ denotes the aggregate tax that worker i is supposed to pay for all other group members.[16] $Envy_{ijt}^{k}$ and $Guilt_{ijt}^{k}$ are worker i's preferences for one's (dis)advantage toward co-worker k that are structurally estimated. $\bar{\hat{a}}_{-ijt}$ is the mean of permanent ability of other teammates who belong to the same group with i.[17] Note that all these four variables are orthogonal to the residual term conditional on individual fixed effects. X is a vector of time-variant worker's covariates. $RepInteract_{ijt}$ is the proportion of those who live in the same village in a given group. Our experimental design takes advantage of the coexistence of local and seasonally migrant workers in the same workplace by randomly mixing them into same teams, which makes $RepInteract_{ijt}$ exogenous.[18] $PastHistory_{ijt}$ is the weighted

[16] In this main specification, we use the kinship tax that is imposed on subject's earned income and in which the stake size is medium.

[17] As for individual permanent ability, we adopt the same method with Bandiera et al. (2010). Detailed estimation method and results are available in online appendix.

[18] Seasonal workers who come from different provinces are organized and managed by one of the workers in the their own pool and they usually migrate to and stay in the villages for a short period. Generally, a organization of seasonal workers change the village where they stay for engaging in a rice planting task every seasons and they never come back to the same village where they have stayed once. This is because that they pick up the next season's village based on the referral by the last village's organizer. Additionally, they are not supposed to work together with local workers.

probability that a worker can correctly obtain the information of the state for other teammates, which is defined as the number of rounds a worker has shared with a partner weighted by the time distance from the present: sharing one round in the 10 rounds ago should be less important than sharing one round in the last round. The random assignment of teammates in every rounds produces exogenous variations in this covariate. ρ_i and η_{jt} are individual fixed effects and group-round fixed effects, respectively.

The parameters of interest are ϕ, μ, and γ, which identify the causal effects of social norms on the worker's contract choice. In addition, τ captures the impact of relative superiority of individual ability to other teammates on the contract choice. Those variables are all aggregated into each team level excluding worker i herself. This is a natural way to construct key variables in our setting: fist, all workers are randomly assigned into several teams consisting of four to seven members and this random assignment is executed in every experimental rounds: second, the experimenters allocate paddy fields to each team: third, after being assigned to a team and a paddy field, all workers are asked to decide which contracts they prefer, FW or IPR: fourth, the workers who belong to the same team are supposed to engage in a rice planting task together during all working time: and finally, it is strictly prohibited by the experimenters to work with the persons who belong to different teams and to plant rice seedlings in the paddy field that they are not assigned. These features in the setting mean that they can not observe the behavior of those who belong to different teams and vice versa. Moreover, the information on a worker's contract choice, outcome, and payment based on whichever FW or IPR is confidential to other teams but is shared and common knowledge within a her own team so that her decision on a contract choice and productivity can not be affected by and can not affect other teams' ones. Therefore, social norms that a worker feels toward her teammates rather than toward an entire society will be relevant and effective in a decision-making stage of our field experiments.

5.2.2 Econometric Results

Table 6 denotes the estimation results based on the above equation. All coefficients are robust to various specifications. The increase of one standard deviation in guilt aversion decreases the probability to choose IPR by 12.6%. So the more guilt averse the worker is, the less likely she is to opt for IPR. Also, the increase of 10% in the kinship tax rate decreases the probability by 21.8%. Social pressure to redistribute income might be a major obstacle to choose the more remunerative option. Furthermore, the increase of one standard deviation in enviousness augments the probability for IPR by 15.1%. The envy

emotion motivates workers to reduce the wag gap between her and teammates by letting her opt for IPR. Social preferences affect the worker's behavior prominently in a contract choice.

5.3 Under Which Conditions Do Moral Sentiments Work?

In this subsection, we examine what triggers social norms to be effective and under which conditions they work.

5.3.1 Econometric Specification

There are two econometric models for identifying the mechanisms to explain why specific social norms come into play in the workplace. First, we estimate the equation (17) by splitting a whole sample into 20 sets of sub-samples. To be more precise, after sorting all of observations by the variable $RepInteract_{ijt}$ we split the observations into 20 sub-samples each of which includes the same number of observations such that the first set includes from 0 to 5th percentile, the second set includes 5 to 10th percentile, and the final set includes 95 to 100th percentile. By doing this we demonstrate *non-linear* heterogeneous effects of social norms on a contract choice, which is varied with the possibility of repeated interaction. Second, the same way to create sub-samples is applied for non-linear heterogeneous effects varied with the availability of past history information.

5.3.2 Estimation Results

Figure 4, Figure 5, and Figure 6 show that the estimated coefficients obtained from the above econometric model for guilt aversion, enviousness, and kinship tax, respectively. Interestingly, all of figures demonstrate that there exist thresholds around 70th percentile at which the coefficients start to become statistically significant. Guilt aversion and social reputation tax (enviousness) affect negatively (positively) a contract choice above the threshold. However, these social norms explain nothing below the threshold. Put differently, social norms come into the play in the workplace only if the workers ensure that the interactions with current partners would infinitely be repeated. The workers might recognize that the game is infinite if the possibility of repeated interaction is above 75% which is corresponding to the experimental settings of Bó (2005) and Bó and Fréchette (2011).

The story is quite different when it comes to the non-linear heterogeneous effects by the observability of current partners' past history. Figure 7, Figure 8, and Figure 9 show

the results of the second econometric model described in above subsection. There are two findings. First, in all of estimation results, the coefficients in the first sample (from 0 to 5th percentile) are statistically significant: negative for guilt aversion and kinship tax and positive for enviousness. Note that the actual values of $PastHistory_{ijt}$ for the 0 to 5th percentile are almost zero. Then these facts allow us conclude that the community might be able to sustain the social norms even without the information sharing with respect to the current partners' history. Second, the coefficients are always statistically significant even for other sub-samples and there might not be any threshold unlike the estimation results of the repeated interaction.

In sum, the community enforcement mechanism supports to sustain social norms only if the repeated interaction is ensured. Additionally, the information of the partners' state such as reputation is likely to be transmitted through both a whole community network and localized social connections, which reinforces the community enforcement of social norms.

6 Discussion and Conclusion

The interplay between economic incentives and social norms is the critical issue in economics. This paper develops new field experiments to identify the causal impacts of social norms on worker's self-selection behavior into the different intensity of incentive contracts. In the two-stage field experiment which is carried out in the traditional agrarian labor contract in the Philippines, we randomly organize teams in which workers are required to work together. Next, we offer two options of contracts for workers, i.e. FW and IPR, and they can choose one of them. Thereafter, randomly picked up subjects who choose IPR are converted into FW. Econometric results indicate that the self-selection effect dominates and explains major parts of the productivity gap between FW and IPR. In addition, social norms significantly alter the decision-making by workers. Guilt aversion and kinship taxation discourage workers to choose the remunerative option or IPR, whereas enviousness facilitates them to opt for it. Finally, even without frequent and fixed interactions within communities, community enforcement mechanisms which are founded by infinite repeated interaction and availability of informal reputation system trigger the emergence of social norms among workers. Although the field experiment targets a specific pool of subjects and villages in a certain timing, external validity in this paper might relatively be assured compared to other existing studies due to relying on a larger sample size or observations, focusing on the real labor contract, and analyzing relevant and prevalent

social preferences.

References

Ashraf, N., Berry, J., & Shapiro, J. M. (2010). Can higher prices stimulate product use? evidence from a field experiment in zambia. *The American economic review*, *100*(5), 2383–2413.

Bandiera, O., Barankay, I., & Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 917–962.

Bandiera, O., Barankay, I., & Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, *77*(4), 1047–1094.

Bandiera, O., Barankay, I., & Rasul, I. (2010). Social incentives in the workplace. *The Review of Economic Studies*, *77*(2), 417–458.

Beekman, G., Gatto, M., & Nillesen, E. (2015). Family networks and income hiding: evidence from lab-in-the-field experiments in rural liberia. *Journal of African Economies*, ejv008.

Bellemare, C., Kröger, S., & Van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, *76*(4), 815–839.

Bó, P. D. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *The American Economic Review*, *95*(5), 1591–1604.

Bó, P. D., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *The American Economic Review*, *101*(1), 411–429.

Bohm, P., Lindén, J., & Sonnegård, J. (1997). Eliciting reservation prices: Becker–degroot–marschak mechanisms vs. markets. *The Economic Journal*, *107*(443), 1079–1089.

Boltz, M., Marazyan, K., & Villar, P. (2015). *Income hiding and informal redistribution: A lab-in-the-field experiment in senegal* (Tech. Rep.). Citeseer.

Cadsby, C. B., Song, F., & Tapon, F. (2007). Sorting and incentive effects of pay for performance: An experimental investigation. *Academy of management journal*, *50*(2), 387–405.

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*, *101*(2), 556–590.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, 817–868.

Festre, A. (2010). Incentives and social norms: A motivation-based economic analysis of social norms. *Journal of Economic Surveys*, *24*(3), 511–538.

Goto, J., Sawada, Y., Aida, T., & Aoyagi, K. (2015). *Incentives and social preferences: Experimental evidence from a seemingly inefficient traditional labor contract* (Tech. Rep.). Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University.

Guiteras, R. P., & Jack, B. K. (2014). *Incentives, selection and productivity in labor markets: Evidence from rural malawi* (Tech. Rep.). National Bureau of Economic Research.

Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, *7*, 24–52.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644–1655.

Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*, *59*(1), 63–80.

Karlan, D., & Zinman, J. (2009). Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica*, *77*(6), 1993–2008.

Kim, H. B., Kim, S., & Kim, T. T. (2017, March). *The Selection and Causal Effects of Work Incentives on Labor Productivity: Evidence from a Two-Stage Randomized Controlled Trial in Malawi* (IZA Discussion Papers No. 10644). Institute for the Study of Labor (IZA). Retrieved from `https://ideas.repec.org/p/iza/izadps/dp10644.html`

Lazear, E. P. (2000). Performance pay and productivity. *The American Economic Review*, *90*(0), 1346–1361.

Leuven, E., Oosterbeek, H., Sonnemans, J., & Van Der Klaauw, B. (2011). Incentives versus sorting in tournaments: Evidence from a field experiment. *Journal of Labor Economics*, *29*(3), 637–658.

Squires, M. (2016). Kinship taxation as a constraint to microenterprise growth: Experimental evidence from kenya.

Umekage, O., Inoue, K., & Watabe, T. (1966). Studies on the simplification of transplanting method of rice plants i: Growth habit of the plants of which seedlings

are horizontally transplanted. *Kyoto Prefectural University Academic Report on Agriculture*, 70–77.

Table 1: Individual productivity by contracts

Contract	Obs	Productivity (meter/10min)	Std.
FW	744	14.74	4.48
IPR	807	22.55	4.93
FW converted from IPR	801	19.41	3.74

Notes: This table shows the level of individual productivity by contracts. Contracts have three categories. The fist one is fixed wage or FW. The second one is individual piece rate or IPR. The third one is FW converted from IPR

Table 2: Parameter estimates of the structural model for inequity aversion

	α_1		β_1		α_2		β_2		λ	P	Offer	γ		Other Parameters	
Constant	-0.478	**	-0.684	**	0.512	**	-0.637	**	-0.771		0	-0.151	*	$V(\varepsilon_{ii}^{P})$	-0.510 **
	(-0.226)		(-0.293)		(0.26)		(-0.278)		(-1.528)			(-0.084)			(-0.265)
Difference in											15	0.716	**	$V(u_{ik}^{\alpha})$	-0.543 **
Sex	-0.662		-0.357		0.961		-0.493		-0.349	0.888		(0.312)			(-0.259)
	(-1.416)		(-0.764)		(2.055)		(-1.054)		(-0.747)	(1.901)	30	0.807	***	$V(u_{ik}^{\beta})$	0.015 **
Education	-0.179		0.076		-0.989		-0.048		0.944	0.261		(0.264)			(0.008)
	(-0.23)		(0.098)		(-1.271)		(-0.062)		(1.213)	(0.335)	45	0.627	***	$V(u_{ik}^{P})$	0.823 **
Age	0.261		0.417		0.593		0.801		0.777	0.486		(0.204)			(0.395)
	(0.432)		(0.688)		(0.979)		(1.323)		(1.284)	(0.802)	55	-0.396	**	$\rho_{\alpha\beta}$	-0.155 *
Income	-0.676	**	0.748	*	0.296	*	-0.991	*	-0.965	-0.106		(-0.167)			(-0.084)
	(-0.339)		(0.395)		(0.16)		(-0.522)		(-0.999)	(-0.11)	70	-0.711	**	$\rho_{\alpha P}$	0.710 *
Sum of												(-0.365)			(0.379)
Sex	-0.762		0.304		0.969		0.025		0.137	0.333	85	-0.768	**	$\rho_{P\beta}$	-0.373
	(-0.916)		(0.365)		(1.164)		(0.03)		(0.165)	(0.4)		(-0.358)			(-0.926)
Education	-0.799	*	0.272		0.612		-0.370		-0.565	0.415	100	-0.785	***		
	(-0.428)		(0.326)		(0.735)		(-0.445)		(-0.679)	(0.499)		(-0.255)			
Age	-0.639	*	-0.096	*	-0.033		-0.531		0.087	-0.134					
	(-0.354)		(-0.051)		(-0.588)		(-0.445)		(0.073)	(-0.112)					
Income	-0.709	*	-0.992	*	-0.571	*	0.454	**	-0.770	0.204					
	(-0.377)		(-0.523)		(-0.31)		(0.229)		(-0.653)	(0.173)					
Both recognized kinship	-0.073		-0.556	*	0.540		0.154	*	-0.249	0.251 *					
	(-0.089)		(-0.297)		(0.657)		-0.229		(-0.303)	-0.173					
Both recognized friendship	0.957	*	0.848		0.950	*	0.456		0.157	-0.603 *					
	(0.516)		(1.865)		(0.502)		(1.003)		(0.345)	(-0.33)					
Geographical distances	-0.963	**	-0.641	*	0.620	*	0.076	**	-0.842	0.460 **					
	(-0.471)		(-0.346)		(0.336)		(0.039)		(-0.971)	(0.196)					

Notes: Standard errors are given in parentheses. *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

Table 3: Descriptive statistics of social reputation tax rate

Size of endowments	Income source	All sample					Kinship tax rate >0				
		Obs.	Mean	Std	Min	Max	Obs.	Mean	Std	Min	Max
Small	Earned	7212	1.58	2.34	0.00	10.92	3833	2.97	2.48	0.03	10.92
(40PHP)	Windfall	7212	2.30	3.29	0.00	14.50	4299	3.86	3.48	0.03	14.50
Medium	Earned	7212	1.50	2.23	0.00	10.49	3808	2.83	2.37	0.01	10.49
(120PHP)	Windfall	7212	2.53	3.30	0.00	13.47	4296	4.25	3.32	0.01	13.47
Large	Earned	7212	1.99	2.88	0.00	13.25	4238	3.74	3.01	0.01	13.27
(360PHP)	Windfall	7212	2.60	3.35	0.00	14.45	4578	4.36	3.33	0.01	14.45

Notes: This is the distribution of social reputation tax rate by size of endowment and income sources. The columns for "All sample" include all observations while the columns for "social reputation tax>0" include only observations for which tax rate is more than zero. A unit of figure is percentage.

Table 4: Self-selection effect vs. incentive effect

Dep Var: Individual Productivity (meter per 10 min)		
	Baseline	First choice
Self-selection effect	4.6712***	5.2837*
(Dummy for taking one if a subject chose IPR)	(0.8464)	(2.6112)
Incentive effect	3.1104***	3.10293***
(Dummy for taking one if a subject worked under FW)	(0.6012)	(0.6899)
Obs.	2352	141

Notes: Standard errors are given in parentheses. *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

Table 5: De-motivation effects by random conversion

Dep Var: Individual Productivity	(1)	(2)	(3)	(4)	(5)
Self-selection effect	4.6712***	4.3020***	4.2892***	4.2293***	4.3933***
(Dummy for taking one if a subject chose IPR)	(0.8464)	(1.0148)	(1.0222)	(1.2922)	(1.3284)
Incentive effect	3.1104***	3.1104***	3.1104***	3.1104***	3.1104***
(Dummy for taking one if a subject worked under FW)	(0.6012)	(0.6012)	(0.6012)	(0.6012)	(0.6012)
First conversion dummy * self-slection effect		0.0284			
		(1.4928)			
Conversion rate = 0.3 * self-selection effect			0.0841		
			(1.3398)		
1{others are not converted} * self-selection effect				0.1012	
				(1.3485)	
1{Incentive effect>average} * self-selection effect					0.2928
					(1.3837)
Group and Round FE	Yes	Yes	Yes	Yes	Yes
Obs.	2352	2352	2977	2352	2352

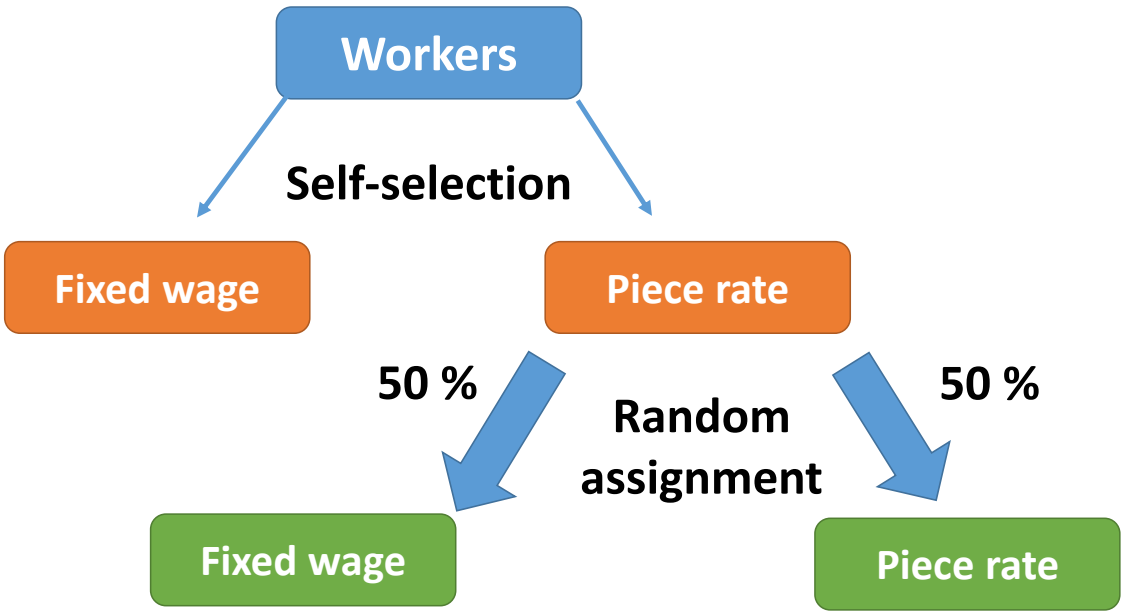
Notes: Standard errors are given in parentheses. *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

Table 6: The determinants of self-selection process

Dep Var: Choosing Individual Piece Rate = 1	(1)	(2)	(3)	(4)
Guilt aversion	-1.4356***	-1.4351***	-1.4372***	-1.4839***
	(0.1054)	(0.1023)	(0.1837)	(0.1384)
Enviousness	1.5664**	1.5632**	1.9021**	1.9831**
	(0.7285)	(0.7285)	(0.8282)	(0.8384)
Kinship tax rate	-2.1754**	-2.1744**	-2.8372**	-2.8032**
	(1.0726)	(1.0722)	(1.4193)	(1.4492)
Other time-variant worker's controls	No	No	No	Yes
Individual FE	No	No	Yes	Yes
Group and Round FE	No	Yes	Yes	Yes
Obs.	2352	2352	2352	2352

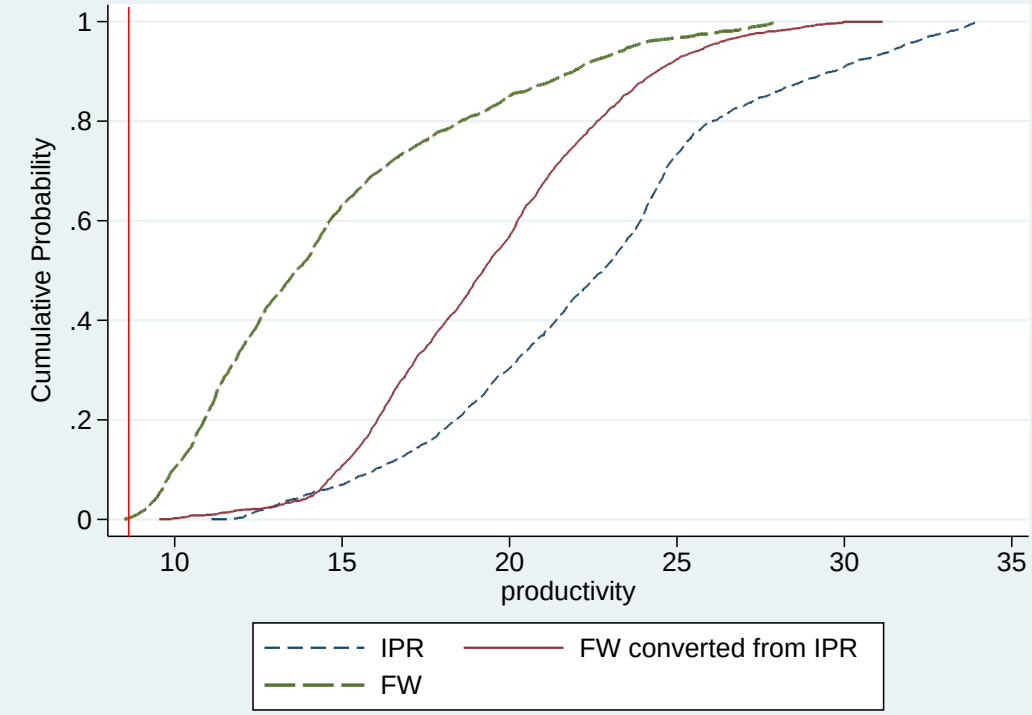
Notes: Standard errors are given in parentheses. *, **, and *** indicate significant at the 10%, 5%, and 1% levels, respectively.

Figure 1: The structure of the two-stage field experiment



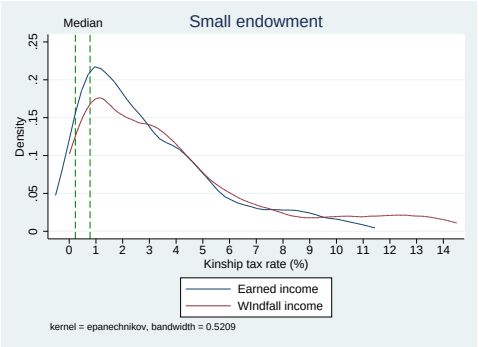
Note: This figure shows the structure of two-stage field experiment. In the first stage, the workers are proposed two contracts or FW and IPR. They are asked to self-select themselves into either contract. The second stage forcefully and randomly converts some of the workers who choose IPR into FW with 50% probability.

Figure 2: The distribution of individual productivity

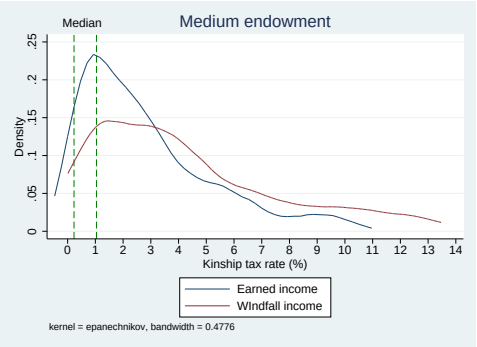


Note: This figure shows that the distribution of individual productivity by contracts.The curves are the cumulative distribution functions. The vertical red line denotes the minimum requirement effort level calculated from the conceptual framework and the field experimental setting.

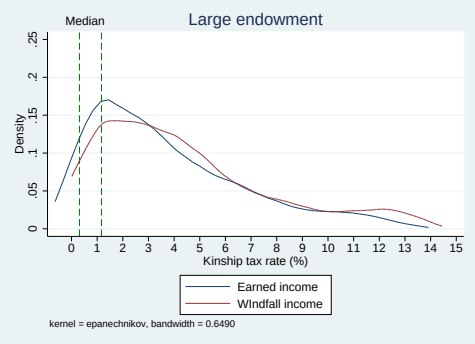
Figure 3: The distribution of social reputation tax rate



Panel A



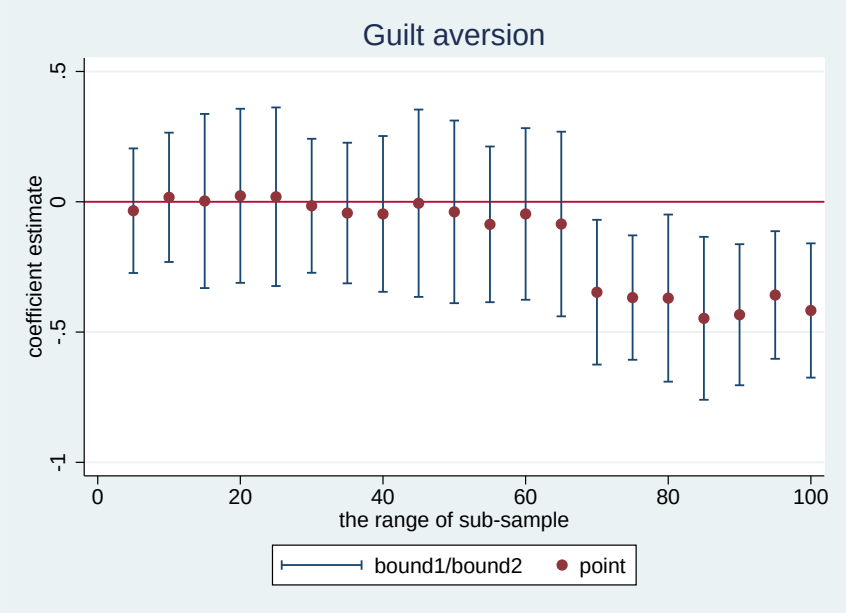
Panel B



Panel C

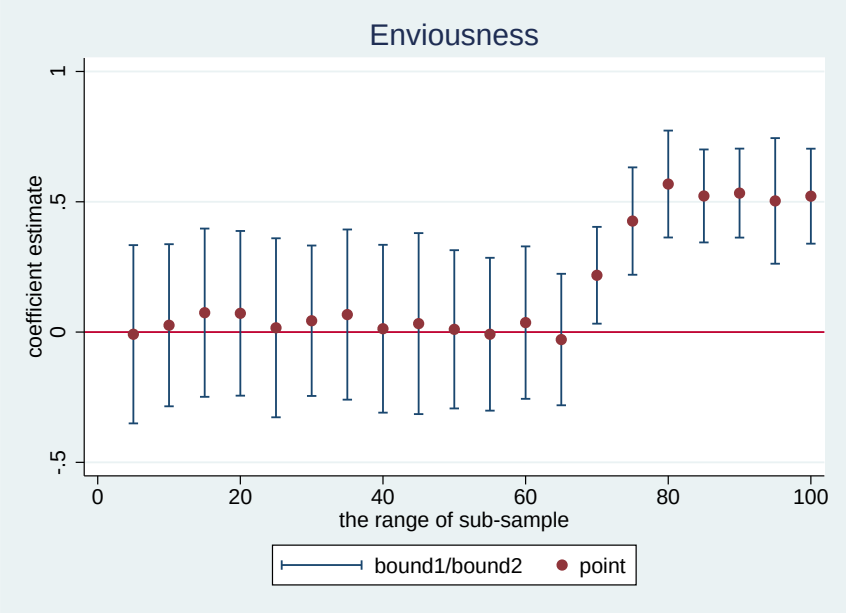
Note: Panel A, B, and C show the distribution of social reputation tax rate for small, medium, and large endowment, respectively. vertical lines colored by green denote median for each distribution.

Figure 4: The coefficients of guilt aversion by the possibility of repeated interaction



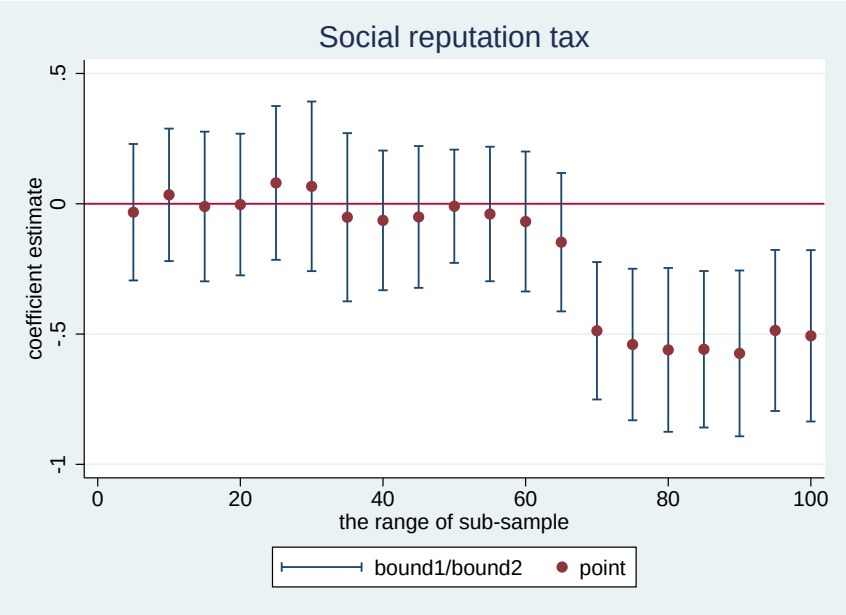
Note: This figure shows the estimated coefficients for guilt aversion by sub-samples. After sorting the observations by the value of the possibility of repeated interaction we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.

Figure 5: The coefficients of enviousness by the possibility of repeated interaction



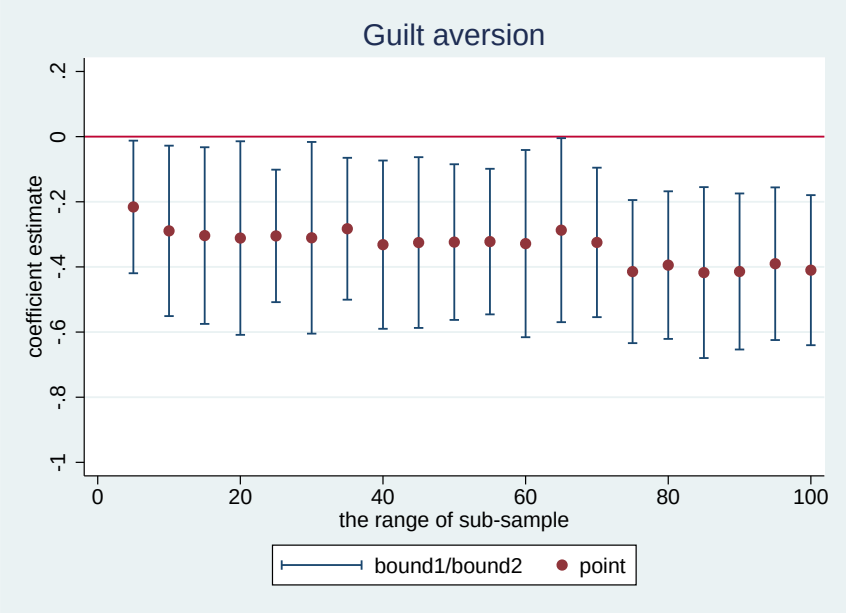
Note: This figure shows the estimated coefficients for enviousness by sub-samples. After sorting the observations by the value of the possibility of repeated interaction we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.

Figure 6: The coefficients of social reputation tax by the possibility of repeated interaction



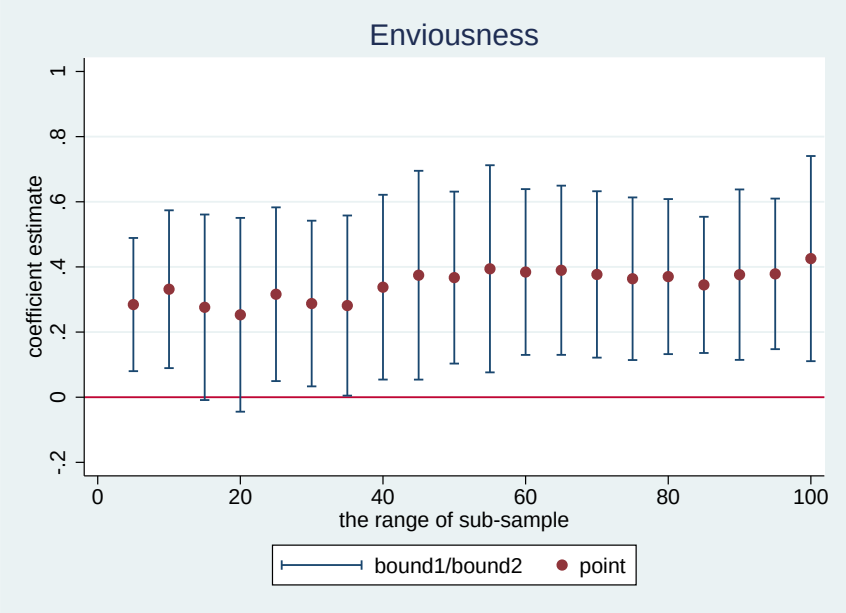
Note: This figure shows the estimated coefficients for social reputation tax by sub-samples. After sorting the observations by the value of the possibility of repeated interaction we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.

Figure 7: The coefficients of guilt aversion by the availability of past history of partners



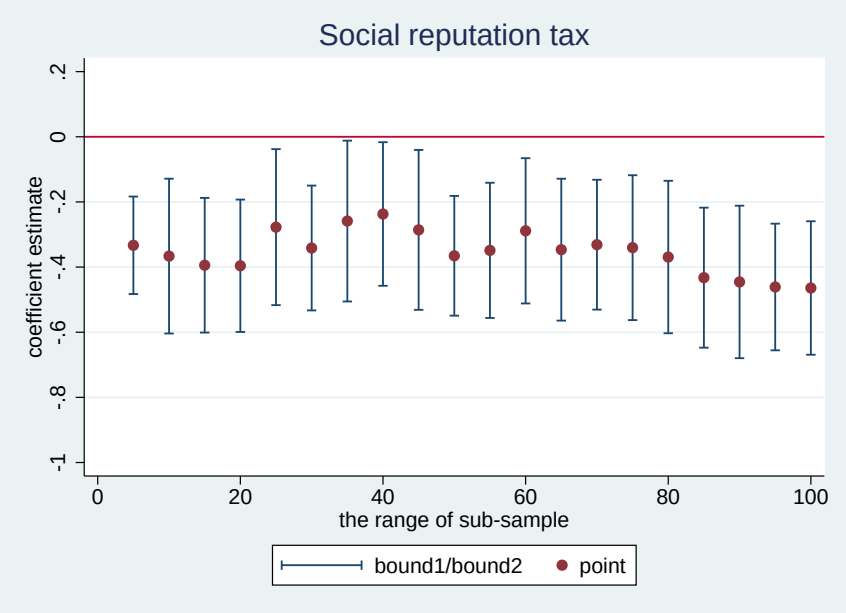
Note: This figure shows the estimated coefficients for guilt aversion by sub-samples. After sorting the observations by the value of the availability of past history of partners we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.

Figure 8: The coefficients of enviousness by the availability of past history of partners



Note: This figure shows the estimated coefficients for enviousness by sub-samples. After sorting the observations by the value of the availability of past history of partners we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.

Figure 9: The coefficients of social reputation tax by the availability of past history of partners



Note: This figure shows the estimated coefficients for social reputation tax by sub-samples. After sorting the observations by the value of the availability of past history of partners we split a whole observation into 20 sub-samples (0-5%, 5-10%, ..., 95-100% of a whole observation) in which the same number of observations is included. The range of coefficients denotes the confidential intervals and point estimates are also plotted.