CSAE Working Paper WPS/2015-08

When No Bad Deed Goes Punished: A Relational Contracting Experiment in Ghana

Elwyn Davies* and Marcel Fafchamps**

*University of Oxford
**Stanford University[1]

March 2015

Abstract

This paper uses experimental methods to study the impact of limited enforcement and reputation on employer-worker relations in labour markets in Ghana. Participants, students recruited from universities in Accra, Ghana are designated as either employers or workers and play a gift-exchange game on a tablet computer. In this game, employers make wage offers to workers, who can then choose to accept or reject and, after accepting, what effort level to exert. Five treatments were used to assess the impact of limited enforcement, competition between employers and reputation. Each participants plays four games, consisting of five trading periods. We find different results from earlier experiments in developed countries: while these experiments have found strong evidence for relational contracting and conditional reciprocity, we do not find evidence for this. We find that a subgroup of workers exerts very low effort levels, but that this low effort of the workers is not punished by employers, who are not responsive in their wage offers to what the workers did previously. As a result, on average, the workers capture most of the profits. Introducing competition or a multilateral reputation mechanism does not significantly improve this.

Keywords Relational contracting, conditional reciprocity, gift-exchange game, punishment strategies, Ghana.

JEL classification C71, D2, D86, E24, O16

[1]Contact details: elwyn.davies@economics.ox.ac.uk (http://www.elwyndavies.nl/) and fafchamp@stanford.edu. These experimental sessions are part of a wider study on Ghanaian entrepreneurship. This document is an output from research funding by the UK Department for International Development (DFID) as part of the iiG, a research programme to study how to improve institutions for pro-poor growth in Africa and South-Asia. The views expressed are not necessarily those of DFID. Elwyn Davies gratefully acknowledges funding from the Economic and Social Research Council (ESRC). Many thanks to Václav Těhle for excellent research assistance, and also thanks to Andrew Kerr and Denise Gray for their assistance during trial sessions of this experiment. Furthermore, we would like to thank Moses Awoonor-Williams for his role in helping out with the logistical matters, as well as our local team in the field, in particular Eric Agyekum and Bismark Owusu Tetteh Nortey. We would like to thank Johannes Abeler, Douglas Bernheim, Stefano Caria, Fola Malomo, Muriel Niederle, Simon Quinn, Hannah Uckat as well as participants of the SEEDEC Conference (Bergen, 2013), the Oxford Firms & Development Research Group (2014) and the Stanford Behavioral and Experimental Lunch (2015) for helpful comments and clarifying questions.

1 Introduction

Economic interactions can be risky when enforcement is lacking. If a supplier does not deliver the right quality of a product, this can have costly adverse effects to the buyer. Ensuring the right quality of a product can be challenging, because legal enforcement mechanisms are often imperfect, especially in developing countries. Besides that, even if enforcement mechanisms are present, it might difficult to objectively prove that the quality was lower than demanded. Empirical studies have shown that labour management is often problematic in sub-Saharan Africa and other developing countries (Bloom & Van Reenen, 2007; Fafchamps & Söderbom, 2006).

In this paper we assess the effect of two enforcement mechanisms, *relational contracting* and *multilateral reputation*, in labour markets in Ghana. For this we use a gift-exchange game experimental design modelled after Brown et al. (2004), which mimics labour market relations using a principal agent setting: employers make offers to the workers, promising a specific wage in return for a specific level of effort, and the workers, after accepting these offers, can choose an effort level to exert. The first enforcement mechanism, relational contracting, relies on repeated interactions between individuals, who both gain enough value out of this relationship that it is not in their interest to have the relationship severed or its value reduced. The threat of termination or decreasing its value can provide an incentive to not cheat. Similarly, a multilateral reputation mechanism, in which information about the trustworthiness of someone is shared, can deter cheating (Milgrom et al., 1990; Greif, 2000).

Our experiments confirm earlier findings that introducing imperfect enforcement leads to a decrease in the level of effort chosen by the workers. Also our experiments provide evidence for conditional reciprocity on behalf of the worker: there is a positive relationship between the wage offered and the effort level exerted by the worker. This is in agreement with earlier experiments on "gift exchange" (Fehr et al., 1993, 1998; Fehr & Falk, 1999; Brandts & Charness, 2004). However, in the treatments allowing for relational contracting, we find no evidence for conditional reciprocity *on behalf of the employers*: despite workers choosing low effort, employers keep offering high wages to these workers, leading to low and often negative payoffs for employers. Many models of relational contracting predict that a low effort level is punished by either terminating the relationship or by lowering the wage (see e.g. Shapiro & Stiglitz, 1984; Kranton, 1996; Ghosh & Ray, 1996). Relational contracting therefore fails as a disciplining device. The reputation mechanism fails as well: we do not find an improvement in effort exerted by the workers when a reputation mechanism in the form of information sharing is introduced.

Our result on the failure of conditional reciprocity on behalf of the employer is a stark contrast with earlier experiments on relational contracting in developed countries. In Brown et al. (2004) students from Zürich (Switzerland) play a gift-exchange game and it is shown that a pattern of relational contracting emerges: long term relationships between workers and employers where high effort provision

by the worker can be sustained, because deviating behaviour is punished by the employers by terminating the contract. In a follow-up paper it is shown that this even holds when there is an excess demand for labour (Brown et al., 2012). In both papers there is conditional reciprocity: a lower choice of effort is followed by either a termination of the contract or a lowering of the wage.

Our paper fits in an emerging literature on how preferences and strategic behaviour differ across societies and cultures (Henrich et al., 2006; Cardenas & Carpenter, 2008; Henrich et al., 2010). By framing this explicitly in a labour market context, we intend to capture heuristics influenced by prevailing social norms and expectations. Our subjects are students from Ghanaian universities, many of whom are business or economic students who expect to either become wage employed. Our results indicate that employing effective labour management is problematic, which is in agreement with earlier empirical findings on labour markets and labour management in developing countries (Bloom et al., 2014; Fafchamps & Söderbom, 2006).

The paper is structured as follows: Section 2 presents the experiment and gives an overview of past experiments and behaviour that we might expect on the basis of that. Several testable predictions are formulated. Section 3 discusses the practical implementation of the experiments in Ghana. Section 4 presents the results and tests the predictions. Section 5 concludes and discusses possible explanations for the behaviour found in this experiment.

2 The experiment

2.1 The gift-exchange game

The experiment is an adaptation of the gift exchange game, adapted to a labour market context. In the game employers make offers to workers, specifying how much they will pay them and the level of effort they desire. The worker then chooses which offer to accept and subsequently which level of effort to exert. A higher level of effort is more costly to the worker, but gives the employer a higher payoff. The worker receives the payment regardless of the effort level chosen. This game is played for five periods, after which the employers and workers are rematched.

The participants are divided in three groups of six. In each group three participants are given the role of employer and three participants are given the role of worker. This allocation is random. Participants can recognize workers and employers using a random identification letter, which is randomized at the beginning of each game and stays the same throughout the five periods of the game. Each period in a treatment consists of three stages:

- **The contracting stage.** In this stage employers and workers contract with each other, in a virtual marketplace. Each of the workers is listed with their identification number clearly visible. This

Effort level e	low (e_L)	medium (e_M)	high (e_H)
Benefit to employer $\Pi(e)$	5	20	40
Cost to worker $c(e)$	0	2	6
Surplus $S(e) = \Pi(e) - c(e)$	5	18	34

Table 1: The payoff structure in the experiment.

stage consists of three steps:

- First, the employer can make offers to every worker individually. An offer specifies the payment that the employer will make to the worker w and the desired level of effort e from the worker.
- Second, the employer observes all the offers that are made by other employers and has then a chance to revise the offers to the workers. The initials offers are not shown to the workers.
- Third, in a randomly across workers determined sequence, each of the workers gets to see the offers made to him/her and then can choose to accept one of the offers. In case the worker rejects all offers or no offers are available, the worker does not receive any payment.

- **The supply stage.** In this stage, the worker first receives the payment w from the employer and then chooses the amount of effort e to exert. This effort can either be low, medium or high ($e \in \{e_L, e_M, e_H\}$) and costs the worker $c(e)$. This level of effort gives the employer a benefit of $\Pi(e)$. The payoffs to the employer π_E and to the worker π_W are:

$$\pi_E = \Pi(e) - w \tag{1}$$

$$\pi_W = w - c(e) \tag{2}$$

The values of $c(e)$ and $\Pi(e)$ for the three different levels of effort are described in Table 1.[2] The choice of effort is not revealed to the employer until the end of the next stage.

- **The rehiring stage.** Before moving to the next period, the employer is given the opportunity to make a direct offer to the contracted worker for the next period. At this stage the employer does not know the chosen level of effort by the worker yet. The employer is presented with the three possible choices of effort by the worker and is asked whether he or she wants to rehire this worker in the next period and for what kind of contract (specifying a wage w and effort level e, as before). Similarly, the worker is asked what the *minimum* payment needs to be in order to accept the offer. If the offer of the employer is equal or higher to this minimum payment, the employer and worker

[2]The payoff structure is designed in such a way that the marginal cost of effort is strictly increasing. The effort levels are equal to three out of the ten effort levels in Brown et al. (2004). The number has been limited to three to allow the strategy elicitation in the rehiring stage and to simplify the game.

Treatment	Number of participants	Enforced compliance	Provision of multilateral compliance information
(1C)	1 person	full compliance	n/a (only 1 partner)
(1E)	1 person	effort choice	n/a (only 1 partner)
(3C)	3 person	full compliance	n/a (compliance enforced)
(3E)	3 person	effort choice	only own information
(3ES)	3 person	effort choice	full information

Table 2: Overview of the five treatments.

will contract with each other in the next period, and proceed directly to the supply stage of the next period.

The *high* level of effort costs 6 points and gives the employer a benefit of 40 points. This is a desirable outcome for the employer, but it requires a high level of trust to coordinate on this outcome: the employer must give the worker a high enough wage so that the 6 points effort cost can be overcome, but this can be risky, because the worker can choose to pick a lower level of effort, which means that the employer might lose out.[3] The *medium* level of effort costs the worker 2 points and gives the employer a benefit of 20 points. This effort choice puts a relatively small burden on the worker, and therefore requires less trust on behalf of the employer to coordinate on this outcome.[4] Finally, the *low* level of effort does not cost the worker any points, but gives the employer a benefit of only 5 points.

2.2 Treatments

The experiment consisted of five different treatments, each with a modified version of the gift-exchange game. Each participant played four times five periods and after each five periods participants were randomized between groups. The treatments differ in three dimensions:

- **Number of market participants.** In some treatments markets only had one employer and one worker, while in other treatments markets consisted of three employers and workers.
- **Enforced compliance.** In some treatments the worker did not have a choice of effort after accepting the offer and had to choose the level of effort that was demanded by the employer, while in other treatments the choice of effort is discretionary after accepting the offer.
- **Information on past actions.** In some treatments an employer only had access to information on compliance by the workers when they were hired by him- or herself, while in other treatments the employer also got information on compliance when the worker was hired by the other employers.

[3] In Brown et al. (2004) this corresponds to the average level of effort in case bilateral relations were allowed (the "ICR" treatment).

[4] In Brown et al. (2004) this was the average level of effort in the anonymous case where identification numbers were randomized after each round, so no (individual) reputations could emerge (the "ICF" treatment).

Sequence	Game 1 5 periods	Game 2 5 periods	Game 3 5 periods	Game 4 5 periods	No. of participants
I	(1C)	(1C)	(1C)	(1C)	28
II	(1C)	(1E)	(1E)	(1E)	48
III	(1C)	(1E)	(3E)	(3E)	36
IV	(1C)	(1E)	(3E)	(3ES)	54
V	(1C)	(3C)	(3C)	(3C)	30
VI	(1C)	(3C)	(3E)	(3E)	54
VII	(1C)	(3C)	(3E)	(3ES)	54

Table 3: The seven treatment sequences of the experiment. After each game of five periods employers and workers were re-matched randomly.

Table 2 shows an overview of the treatment variations used in the experiment. Seven different sequences of treatments were used, which are shown in Table 3. The sequencing has been set up in such a way that more complicated versions of the game follows simpler versions of the game. To allow for between-group comparisons, some participants were assigned a sequence in which they kept playing the treatments with the simpler version of the game. The first treatment for each participant was always treatment (1C), in which every employer has only one worker and effort is strictly enforced. This treatment is essentially an adapted version of the ultimatum game, because the worker can only choose to accept or reject the proposed split.

Treatments (1E) and (3C) are versions of the game that are half-way between the ultimatum game and the normal gift-exchange game. In treatment (1E) each employer has still only one worker to make offers to, but the worker is free in selecting the level of effort, while in treatment (3C) effort is strictly enforced, but the employer has now three workers to make offers to and each worker has three employers to accept offers from (essentially a multiplayer ultimatum game).

Treatment (3E) is the standard gift-exchange game, where employers only have information about their own worker's past compliance. When making offers they get a graphic overview of past compliance of each of their past workers. This information can also be requested by requesting a history screen, which keeps track of past interactions within each treatment. Only bilateral reputation mechanisms can emerge here. In treatment (3ES) the possibility of multilateral reputation mechanisms is introduced: information on past compliance is available to all employers, by showing a graphical overview of past compliance of the workers when they worked for other employers. Workers are explicitly told at the beginning of the treatment that information on compliance is shared with all employers.

Our experiment differs from Brown et al. (2004) in the following ways: first of all, the hiring process is not a real-time interaction between workers and employers, but takes place in stages. This allows for an easier implementation in a mobile laboratory and reduces variability of responses due to differences of subjects in their technical ability to work with tablets. Second, a rehiring stage is introduced to gain a better insight in contingent contract renewal by employers. Third, we have reduced the numbers

of level of effort to three levels (high, medium and low), which allows for a simpler application of the strategy method in the "rehiring" stage. Fourth, the number of workers and employers is equal and set to one or three depending of the treatment. The (1C) and (1E) are explicitly introduced to rule out competition effects, such as overbidding by employers (winner's curses). [5] Fifth, we reduced the number of periods of each interaction, and introduced a sequential treatment design, where each participant plays four games of five periods in a row. This facilitates understanding of the game, as participants always start with playing the simplest treatment of the game (treatment (1C), which lacks effort choice on behalf of the worker). Finally, our experiment is explicitly phrased in labour market terms, instead of describing work as a "good", an employer as a "buyer" and a worker as a "seller". Our pilot tests showed that using more neutral terms decreased the understanding of the game.[6]

The gift-exchange game is closely related to the trust game (see Berg et al., 1994; Camerer & Weigelt, 1988). The crucial difference is that in the gift-exchange game the size of the surplus is determined by the choice of effort by the worker (the receiver), while in the case of the trust game the size of the surplus is determined by the size of the transfer payment made by the proposer of the offer. This game has been used extensively to study trust (Glaeser et al., 2000; Karlan, 2005; Bohnet et al., 2010). Also in the gift-exchange game trust plays an important role: an employer can only incite a worker to choose high effort when the amount of compensation is high enough. However, a high amount of compensation means that the employer makes himself or herself vulnerable to the possibility that the worker chooses a lower level of effort than demanded, which could lead to high losses for the employer.

2.3 Theory

The gift-exchange game is a sequential prisoner's dilemma game, where the employer makes a first move, and the worker can then respond. In the case of a prisoner's dilemma game, where players can choose to cooperate and defect, it is well established that a cooperative outcome could emerge. The Folk Theorem shows that a Nash equilibrium with any payoff above the minmax payoff is feasible if the players conduct a grim trigger strategy. In a grim trigger strategy, a player will defect for all remaining periods as soon as the other player defects. Axelrod (1984) shows in a repeated prisoner's dilemma game that a very similar tit-for-tat strategy, where one player always copies the action of the previous player (so cooperate if the other player cooperated in the previous round and defect if the other player defected), is optimal against a wide range of strategies employed by other players. The grim trigger strategy and tit-for-tat strategy differ to the extent that redemption is possible: in tit-for-

[5] The equal number of workers and employers is expected to affect the division of the surplus between workers and surplus. In Brown et al. (2004) there are 7 employers and 10 workers, meaning that there is excess supply of labour. This gives employers a slight edge over the workers. Brown et al. (2012) show in a follow-up experiment that changing this ratio of workers and employers to 10 employers and 7 workers affect the division of surplus, but not affect the main efficiency results.

[6] Especially the role of the buyer and seller confused participants: as it is the buyer (i.e., the employer) who specifies the price of the product, while normally in transactions the price is often determined by the seller (i.e., the worker). To avoid the strongest framing effects, we did try to avoid the word *wage* and used *amount* instead.

tat this is possible, but in the grim trigger strategy this is ruled out: as soon as one player defects, the other player will defect forever after this. However, on the equilibrium path both strategies will lead to cooperation, given that the other player is cooperating as well.

However, the game played here is a finite period game. If we assume that workers only care about maximizing their own payoffs each period and are perfectly informed about the other person's payoff, we can use backwards induction to predict the subgame perfect equilibrium: in the last period workers will choose low effort, regardless of the wage offered by the employer. Anticipating this, employers will choose a minimum wage of zero, such that the worker is indifferent between accepting or rejecting the offer. By backwards induction, the subgame perfect equilibrium is that the employer offers a wage of zero points in every period, regardless of past behaviour of the worker, and that the worker chooses low effort in every round, regardless of the amount offered by the employer: cooperation unravels.

Several theories have been proposed to rationalize cooperation in finite period games. Some of these theories assume some sort of irrational behaviour to come to this result: Kreps et al. (1982) show that if a player has a belief that there is a small probability that another player is acting "irrationally" (i.e. the player is not playing subgame perfect equilibrium of not cooperating), and is for example following a tit-for-tat strategy, it is optimal for you to cooperate as well. Fudenberg & Maskin (1986) use a similar argument to show that in a finite repeated game an equilibrium is possible where in the first periods of the game the game is played like an infinitely repeated game, and where players play a cooperation with punishment strategy, while switching to behaviour consistent with backward induction in the last periods of the game.

A similar result as Kreps et al. (1982) can be achieved, not by introducing irrationality, but by incorporating inequity aversion and fairness concerns into the utility function of some of the players. Fehr & Schmidt (1999) present a model where for the utility function does not only depends on the player's own payoff, but also on the difference between payoffs of the two players:

$$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}. \tag{3}$$

In this equation α_i represents the disutility from having a lower payoff than the other player and β_j represents the disutility of having a higher payoff than the other player. Generally, it is assumed that $\alpha_i \geq \beta_i$: a player has a higher disutility from disadvantageous inequality (represented by α_i) than from advantageous inequality (represented by β_i).

Fehr & Schmidt (1999) and Brown et al. (2004) show that even when not all workers have high values of β_i cooperation can emerge in a finite repeated game: if employers respond positively on workers choosing high effort, purely self-interested workers (with $\beta_i = \alpha_i = 0$) have an incentive to mimic the behaviour of "fair" workers (those with higher levels of α_i and β_i) to establish a reputation, because

this behaviour will be reciprocated by higher wages in the next periods.[7] Only in the last periods this imitation does not pay off anymore, and the self-interested workers will choose low effort in the final rounds of the game.

2.4 Predictions

On the basis of the theory, we can formulate a number of hypotheses of how participants play a gift-exchange game as in our experiment. (1) We expect that employers offer high wages to induce high effort by workers, if the combination of a high wage and a high level of effort is profitable to them. (2) Workers to choose high effort as long as the employer offers a high wage. (3) Any deviating behaviour is punished by low effort or a low wage.

We can use the Fehr-Schmidt model of inequality aversion to calculate the optimal response of a worker to a wage offered by the employer. In Figure 2 the optimal response of the worker in a one period version of the gift-exchange game of our experiment is given for three different wage levels (see Figure 1 for the optimal response for all wage levels), and varying parameters of α_i and β_i. For all wage levels the worker will choose low effort when $\alpha_i = \beta_i = 0$, as in the case without inequality aversion. However, for higher levels of β_i, the worker is wiling to choose a higher level of effort. For example, when the wage $w = 19$ a worker will choose high effort for values of $\alpha_i = \beta_i = 0.5$.

The first and second predictions follow from the observation that a fair worker is only willing to choose a high level of effort when the wage is sufficiently high. This is illustrated in Figures 2 and 1: for a given value of α_i and β_i, effort choice is increasing in w. Obviously, the employer will only offer a high wage, if the employer expects that the profit of the combination of a high wage with a high level of effort is more profitable than the combination of a low wage and low effort. The third prediction follows from the threat of punishment after deviations that is supporting the "fair" equilibrium in the Fehr-Schmidt model. On the side of the employers, we expect to contract on the basis of past behaviour of the worker. In particular, a low level of effort is followed by a punishment, either by terminating the relationship or lowering the amount offered. But punishment behaviour could have other sources as well: it could be an emotional response to the behaviour of the worker, for example.

The game is still finite, and therefore we expect some cooperation breaking down in the final periods of the game. We could therefore expect that incomplete contracting in treatments (1E), (3E) and (3ES) reduces high effort provision, especially in the last periods.

[7] This argument is similar to the experimental finding of Camerer & Weigelt (1988) in the setting of an abstract lending game: borrowers whose payoff functions do not favour paying back loans, will actually pay back loans to establish a reputation that they are paying back, in order to encourage the lenders to make a loan.

2.5 Previous experiments

Experimental work has provided evidence for these predictions. Brown et al. (2004) show empirically that in a finite gift-exchange game cooperation is possible. Workers and employers engage in a bilateral relationship based on trust, cooperation and reciprocity. These relationships exhibit high wages and high effort levels for most of the periods, followed by a small drop in the final two periods. The surplus of the transaction is shared approximately equally. Employers use a policy of *contingent contract renewal*, terminating the relationship if the worker does not choose a high level of effort. The threat of economic loss if the relationship is ended makes these bilateral relationships feasible. Follow-up experiments have confirmed these characteristics of successful bilateral relationships (see e.g. Brown et al., 2012; Altmann et al., 2014).

Not only in gift-exchange games, but also in other games it has been established empirically that cooperation is possible in a finite game. Selten & Stoecker (1986) and Andreoni & Miller (1993) show that in a finite period prisoner's dilemma game subjects tend to cooperate for some period of time, until the last couple of periods, when people defect. Other work has put bounds on people's cognition: for example, Johnson et al. (2002) provide experimental evidence that people do not always look ahead in finite period games and that backward induction might be a too strong requirement (even though they do show that the concept is easily taught and learnt).

Treatment (1C) and (3C) are essentially adapted versions of the ultimatum game, in which a "sender" can make a proposal to a "receiver" how to divide a payment between them. The payment will only be made if the "receiver" agrees. In our setting, the employer corresponds to the "sender" and the worker corresponds to the "receiver". In this game, all outcomes of the game where the receiver accepts the offer are efficient. The rational outcome of the game is for the receiver to always accept any offer with a positive transfer payment and for the sender to set the transfer payment as low as possible. The rational outcome is also efficient. However, both lab experiments and field experiments using various subject pools have shown that receivers tend to decline offers with "unfair" distributions (Henrich et al., 2006; Roth et al., 1991; Güth et al., 1983; Güth & Tietz, 1990).

Several experiments have shown that reputation can indeed function as an enforcement mechanism. Image scoring games, which show the history of play, such as Bolton et al. (2005), show that providing reputational information increases the amount that is given to participants that donated more in previous games. Charness et al. (2011) finds that in a trust game people build up a reputation, even if the payoff for doing so is low.

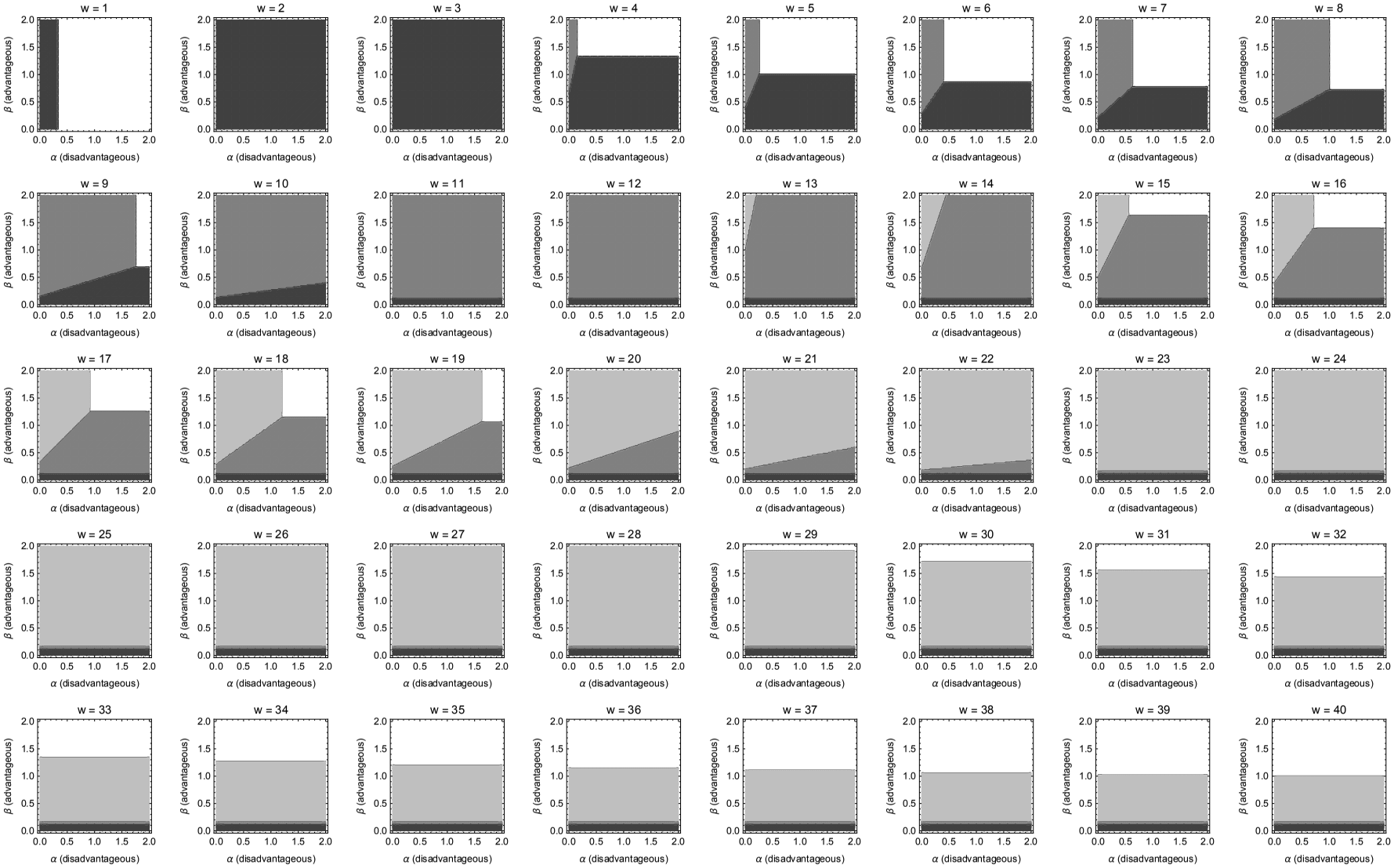


Figure 1: Expected behaviour by worker in the fairness model for different parameters of α_i and β_i. See Figure 2 for the legend of the colours.

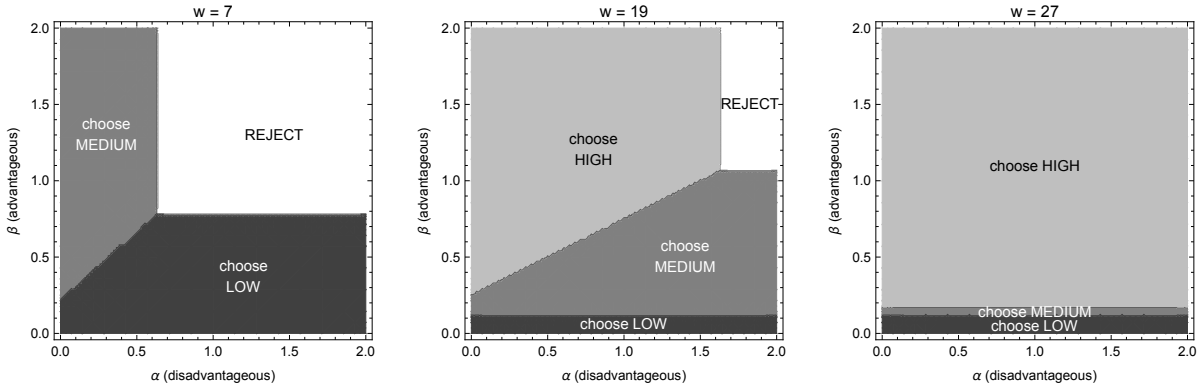


Figure 2: Expected behaviour by worker in the fairness model for different parameters of α_i and β_i, as response to the wage w offered by the employer. Figure 1 shows expected behaviour for all possible values of the wage w.

3 Implementation

For this experiment students from colleges and universities in Accra, Ghana, were recruited. Each session contained 18 to 20 participants and in total sixteen sessions were held. In total 304 participants participated in the study. The points that participants earned during the session were converted to Ghana cedis at the end of the session, with an exchange rate of 0.05 Ghana cedi for every point. Each session lasted between 2 and 2.5 hours. Participants that showed up on time earned a 10 Ghana cedi bonus in addition to their earnings in the experiment. The average earning was 32.2 Ghana cedis, which is about 10 British pounds (about 15 US dollars).

The experiment was conducted using CSAE's Mobile Lab, which is a tablet-based mobile laboratory that can operate completely independently from mains electricity or existing network structures. Each participant was given a 7-inch Android tablet with a custom-built app installed, which was used for user input and communicating with the central server using a wireless connection. Each session started with a 15 minute instruction on how to use the touch screen of the tablet computer, followed by an extensive demonstration of how the game is played. The experiment was completely conducted in English, the common language of instruction in Ghanaian higher education. Visual on-screen aids were used to tell participants of the prospective earnings every time they were about to make a specific choice (such as making an offer), to make sure that participants were aware of the payoff structure of the game.

4 Results

This section describes the results. We start by comparing treatment (1C), which has perfect enforcement, with treatment (1E), which has imperfect enforcement: workers have a choice of which level of effort they would like to exert. In both treatments, markets consist of only one worker and one employer. We

Treatment	Description	Game 1	Game 2	Game 3	Game 4	Total
(1C)	1-person, perfect enforcement	304	28	28	28	388
(1E)	1-person, effort choice	-	138	48	48	234
(3C)	3-person, perfect enforcement	-	138	30	30	198
(3E)	3-person, effort choice (bilateral)	-	-	198	90	288
(3ES)	3-person, effort choice (multilateral)	-	-	-	108	108
Total		304	304	304	304	1216

Table 4: The number of participants in each treatment in each part of the sequence.

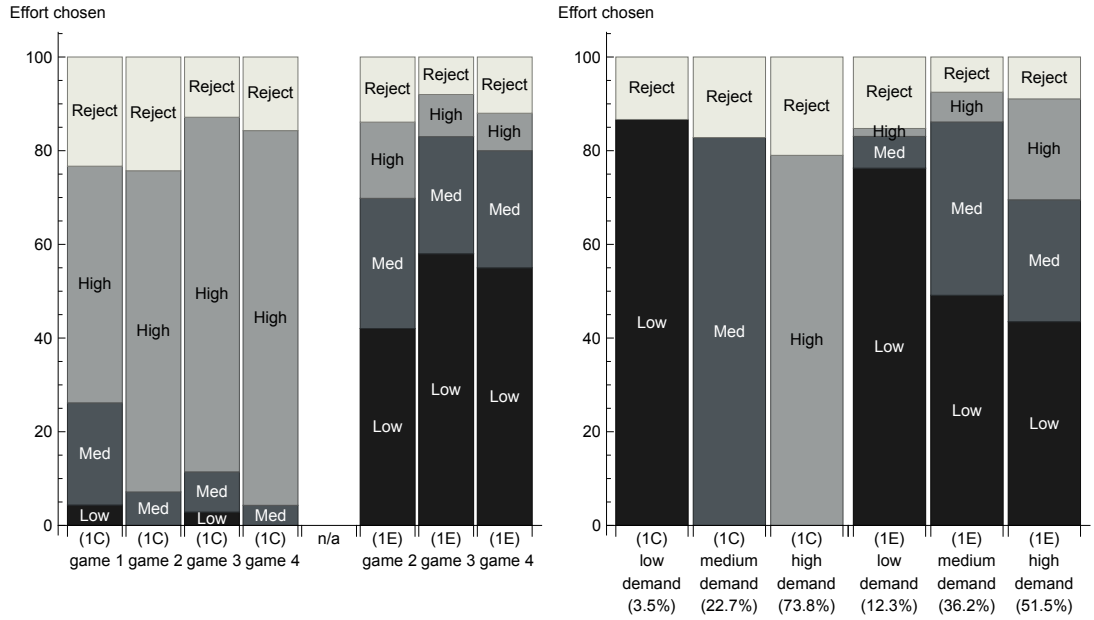


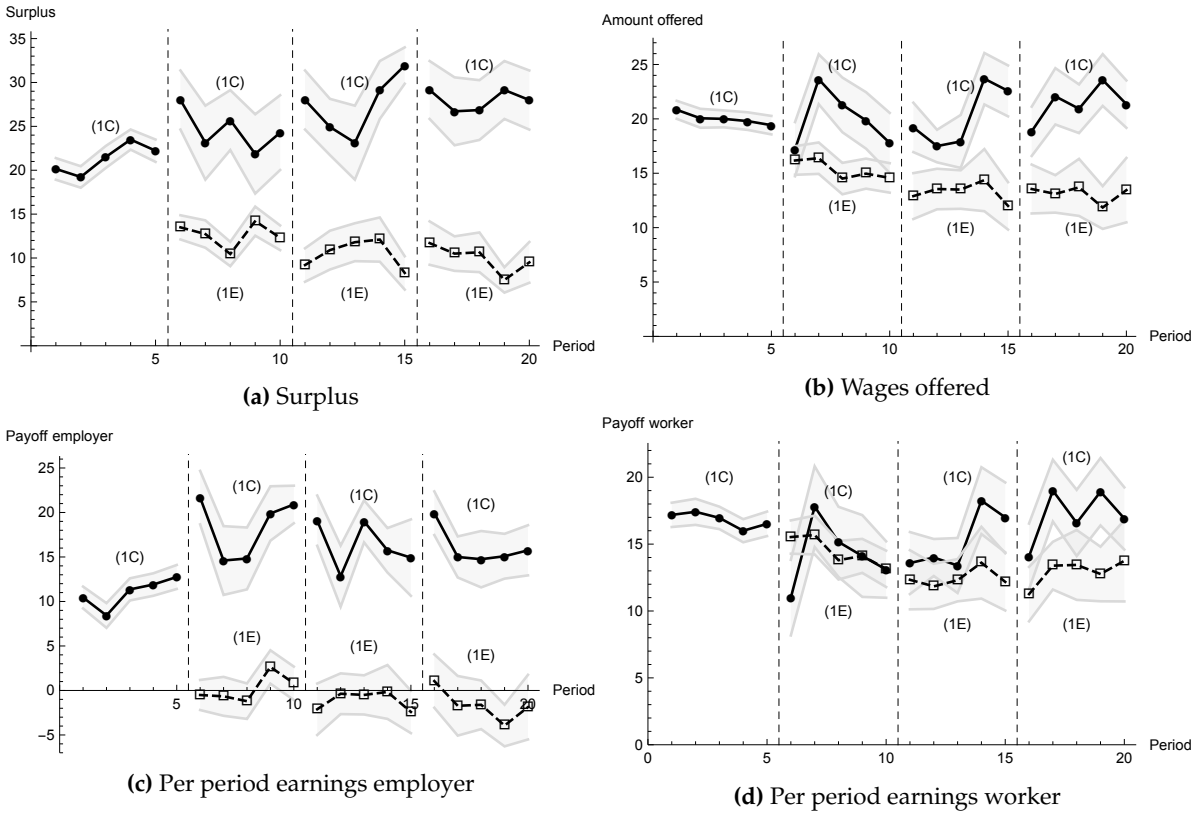
Figure 3: Comparison of the (1C) and (1E) treatment. Left: the choice of the worker in the four games of the perfect enforcement (1C) treatment and the three games of the (1E) treatment, aggregated over all five periods in each game. Right: the choice of the worker following a low, medium and high demand by the employer in the (1C) and the (1E) treatment. The figures are aggregated over all games.

find that imperfect enforcement indeed leads to an average lower effort provision than in the perfect enforcement treatment. We do not find evidence for conditional reciprocity when looking at the patterns of relational contracting.

When comparing the treatments with one worker and one employer with treatments with multiple workers and employers, we do not find significant differences. We do not find strong evidence that reputation functions well as an enforcement device.

4.1 Imperfect enforcement

In the effort choice treatments (1E), (3E) and (3ES) a worker has several options after having received an offer: (1) reject the offer, (2) accept the offer and exert high effort, (3) accept and exert medium effort and (4) accept and exert high offer. In the perfect compliance treatments (1C) and (3C) the worker can only choose to accept or reject the offer and has to comply with the effort demand of the employer after



(a) Surplus

(b) Wages offered

(c) Per period earnings employer

(d) Per period earnings worker

Figure 4: Comparison of treatment (1C) and (1E). The first five periods correspond to game 1, the next five to game 2, and so on. A surplus of 34 points corresponds to high effort, a surplus of 18 points to medium effort and a surplus of 5 points corresponds to low effort. In case no contract is accepted, the surplus is zero points.

accepting the offer: there is perfect enforcement of effort.

Imperfect enforcement in the effort choice treatments does lead to lower effort choices by the workers. In Figure 3 we can see the effort levels exerted by workers in treatment (1C) and (1E). The figure on the left shows the workers' decision in the different sequential games and treatments, while the figure on the right shows the workers' decision as a response to the demand of their employer. In treatment (1C) most of the inefficiency is either caused by the worker rejecting the offer or the employer demanding for a low or medium level of effort. In treatment (1E) we see that a substantial portion of the workers choose low effort: in more than half of the cases (54.6%) the worker chooses low effort in treatment (1E), compared to only 4.5% in treatment (1C). While in treatment (1C) the choice of low effort is purely determined by the employer choosing low effort, as compliance is enforced, in treatment (1E) a large part of the employers (51.5%) demanded high effort.

Imperfect enforcement leads to a lower combined payoff of workers and employers: we find that the surplus, which is the sum of earnings of the worker and employer, is significantly lower in the treatments without perfect enforcement. A fixed effects reduced form regression shows that the surplus is 15.0 points lower in the effort choice treatment (1E) compared to the perfect enforcement treatment (1C). This figure is significant at a 1% level. [8] This is also visible in Figure 4a, which shows the average surplus in the (1C) and (1E) treatments over time. This Figure also shows that in (1C) there is an increase in the amount of surplus over time, while in treatment (1E) the surplus decreases. This is confirmed by the reduced form regressions: for treatment (1C), in game 4 the surplus is 4.7 points higher than in game 1, and for treatment (1E), in game 4 the surplus is 2.5 points lower than in game 2. These figures are significant at a 5% level.

Introducing imperfect enforcement also impacts the wage: a fixed effects reduced form regression shows that the wage drops by drops by 6.3 points in treatment (1E) compared to treatment (1C). This figure is significant at the 5% level. This difference is also visible in Figure 4b, which shows the average wage in the (1C) and (1E) treatments over time. Figures 4c and 4d shows the average per period earnings by the employers and the workers. We see that the earnings of the employers are close to zero or even negative when imperfect enforcement is introduced. The earnings of the workers are at the same level, or a bit lower. Imperfect enforcement seems to impact allocative efficiency: the earnings of the employers drop significantly, while the earnings of the workers drop or remain the same on average.

4.2 Patterns of relational contracting

In Section 2.4 we formulated a number of predictions of cooperation with relational contracting. In this subsection we analyze to what extent we find evidence for these predictions.

[8] These figures relate to fourth game that the participants play, i.e. periods 16-20. See appendix section A.2 for a full comparison.

Wage setting by employers Even though imperfect enforcement leads to a drop in the wage level, the wages offered in the imperfect enforcement treatments, on average between 11.7 and 15.3, are still above the maximum payoff for an employer when a worker chooses low effort, which equals five points. It is also higher than the wage offered of zero points or one point as predicted by a rational model with backward induction. This higher setting of wages is consistent with our prediction that relational contracting leads to higher wages.

Conditional reciprocity by workers As discussed earlier, backwards induction predicts the choice of low effort by a worker. However, previous experiments have shown that there is a positive relation between the offer made by the employer and the response by the worker: the higher the wage offered by the employer, the higher the probability that this higher wage is reciprocated by the worker choosing a higher level of effort (Fehr et al., 1998; Brown et al., 2004, 2012).

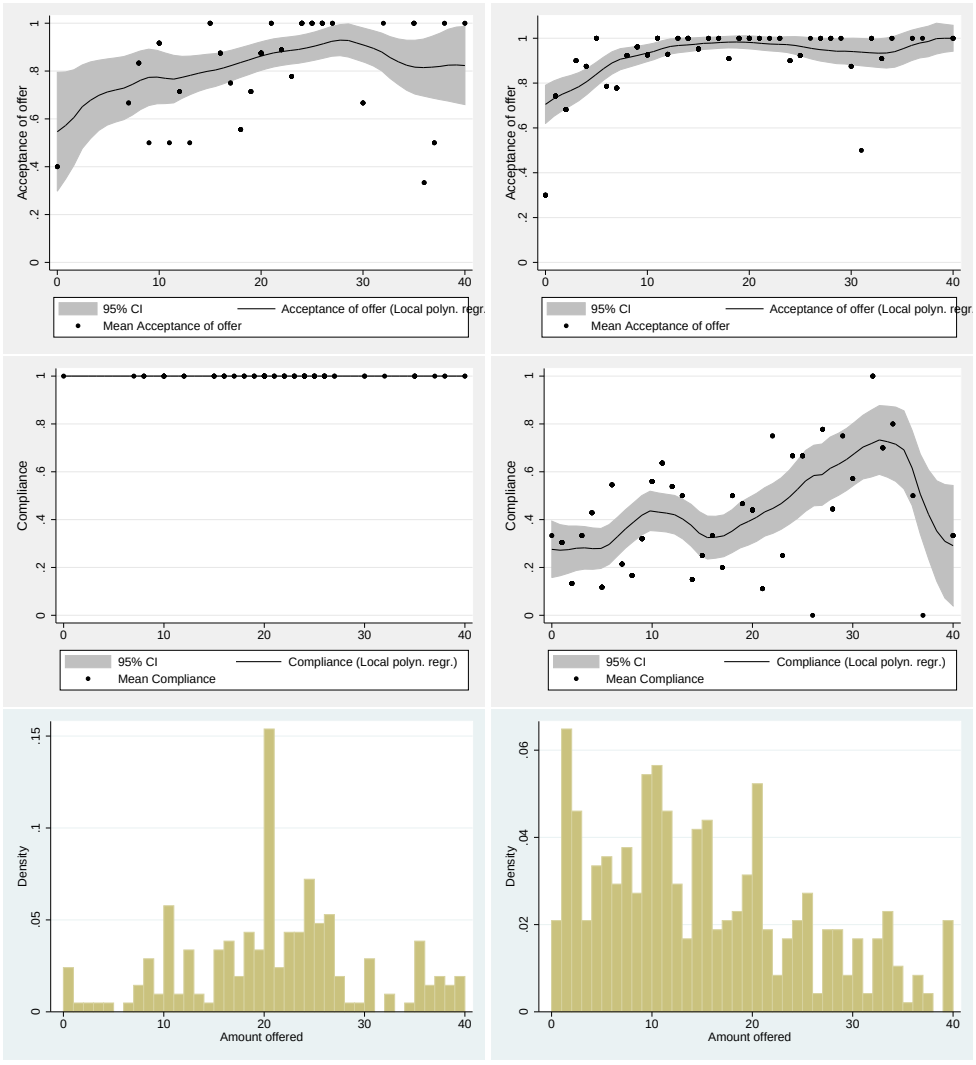
In our experiment we find evidence for a positive relationship between wages and effort levels. Figure 5 shows the relation between the wage and the acceptance of offers, the relation between the wage and the rate of compliance and the distribution of wage offers made for treatment (1C) and (1E). First of all, there is a positive relationship between the wage and the acceptance of the offer in both treatments. Too low offers, which could cause a negative payoff for the worker, are rejected, but also offers where the payoff of the worker is significantly lower than the payoff of the employer are rejected. This is in line with earlier ultimatum game experiments in both developed and developing countries, which show that offers proposing a disadvantageous split for the worker are rejected (see e.g. Cardenas & Carpenter, 2008, for an overview).

Second of all, in the (1E) treatment there is a positive relationship between the wage offered and the effort exerted, with a drop at high levels of the wage.[9] This result is confirmed by the linear probability model regressions in Table 5. In regression (2) it can be seen that an increase of the wage by one point increases the probability of compliance with a high effort offer with 1.56 percent.

Nevertheless, compliance rates seems to be low. The compliance with offers asking for high effort in the bilateral effort choice treatment (1E) is only 23.7 percent on average, and compliance with offers between 18 and 20 points (inclusive), which correspond to more or less equal sharing of surplus, is only 26.7 percent. This low compliance has a dramatic impact on the payoff of the employers in the treatments with effort choice: the average earnings of employers in the (1E) treatment is negative and equal to -0.4 points, compared to 12.5 in the (1C) treatment.[10] Even though the workers on average face lower payoffs in the effort choice treatments, this drop is less dramatic for them than for the employers:

[9]This is mainly driven by low effort choice following offers of 40 points, as there are few offers between 35 and 40. The offers of 40 points are difficult to rationalize: even in the best scenario, when the worker chooses high effort, the payoff to the employer is equal to zero.

[10]To prevent negative earnings at the end of the session all subjects receive an initial endowment, from which negative earnings are substracted. This initial endowment is in addition to the show-up fee, which all participants receive regardless of performance.



(a) Treatment (1C) **(b)** Treatment (1E)

Figure 5: Top: The relation between the wage offered and acceptance. Middle: The relation between the wage and compliance in the (1E) treatment. Compliance means that the worker chooses the effort level specified by the employer, or a higher effort level. Bottom: the distribution of offers.

Dependent variable: Compliance	(1) (all)	(2) (high)	(3) (medium)	(4) (medium & high)
Wage	0.0130**	0.0156**	0.0100	0.0146**
	(0.00346)	(0.00495)	(0.00717)	(0.00375)
HIGH demanded?	-0.963***			-0.312***
	(0.0395)			(0.0679)
MEDIUM demanded?	-0.667***			
	(0.0616)			
Wage × HIGH demanded?				
Wage × MEDIUM demanded?				
Constant	0.969***	-0.0330	0.398	0.291**
	(0.0946)	(0.154)	(0.244)	(0.101)
Observations	434	224	160	384
R-squared	0.583	0.505	0.706	0.528
Adjusted R-sq	0.497	0.297	0.513	0.417
Period dummies	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: The relationship between wage and compliance by the worker in the (1E) treatment. Compliance means that the worker chooses the effort level specified by the employer, or a higher effort level. Column (1) shows the relationship for all wage offers, while columns (2), (3) and (4) show the relationship for all offers where the employer demanded respectively high, medium and medium or high effort. Worker fixed effects are included. Note that for accepted offers demanding low effort, compliance is certain, as any effort choice complies with the effort demanded by the employer. For this reason offers with a low effort demand are excluded in columns (2), (3) and (4).

the average period earnings drop from 16.5 in the full compliance treatment (1C) to 13.4 in the effort choice treatment (1E). In Figure 4c it can be seen that these negative earnings are found in all periods of the game, and do not decrease in later periods.

Conditional reciprocity by employers A central point of many prominent models on cooperation, is that cooperation is sustained by the threat of punishment. In punishment strategies, such as the grim trigger strategy or tit-for-tat, behaviour that is causing harm to the other player is punished. As mentioned earlier, in previous gift-exchange game experiments the employers seem to follow a strategy with contingent contract renewals: high effort choices are rewarded with a high wage in the next period, while low effort choices are punished by terminating the relationship or offering lower wages. This threat of termination fosters cooperation between the workers and the employers.

In our experiment we find low levels of compliance and negative earnings of employers in the game. However, we find little evidence of employers being responsive to low effort exerted by the worker. Table 6 shows the relationship between having not complied in the previous round and the wage offered in the next period. The difference in wage between compliance and non-compliance is 0.283, which is small compared to the difference of 35 points in employer's earnings between high effort and low effort, and this difference is also not significant. We thus find little evidence for punishment behaviour on behalf of the employer in the effort choices treatment.

	(1) Wage offered	(2) Wage offered
Prev. compliance?	0.283 (0.902)	
Prev. rejection of offer?	-1.192 (1.376)	-1.087 (1.339)
Prev. compliance (high)?		0.744 (2.356)
Prev. compliance (medium)?		0.875 (0.945)
Constant	16.79*** (1.239)	16.66*** (1.414)
Observations	380	380

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Only treatment (1E)

4.3 The role of competition and reputation

In treatments (1E) and (1C) markets only consist of one worker and one employer. In treatments (3C), (3E) and (3ES) markets consist of three workers and three employers. Reduced-form regressions show that in treatment (3E) the level of surplus is not significantly higher than in treatment (1E). There is an indication of an effect on the wages offered: wages are 3.0 points higher in treatment (3C) compared to treatment (1C) in game 2 and 2.5 points higher in treatment (3E) compared to treatment (1E) (see also Table 11 in Appendix Section A.2). These increases are significant at the 5% level. The effect of competition seems to be temporary: this significance in increases cannot be found in later games in the sequence; for example, the increase of the wage level in treatment (3C) in game 4 compared to treatment (1C) in game 4 is only 0.1 points and not significant.

The introduction of a multilateral reputation mechanism in treatment (3ES) does not have a significant overall impact on surplus or wage, compared to the treatment with only a bilateral reputation mechanism, treatment (3E). The increase in surplus in treatment (3ES) of 1.0 compared to treatment (3E) is not significant (p-value: 0.539) and neither is the decrease in wage of 0.2 (p-value: 0.801).

In Table 7 we analyze the response of workers to the wage set for the treatments with three workers and three employers, in a similar fashion as in Table 5. The Table presents a linear probability model, with compliance as the dependent variable. Just like in the one-to-one treatments, we find that workers are more likely to comply with higher wage offers. Introducing information sharing is associated with an increase of a 8.1% likelihood of choosing high effort, but this increase is not significant. Information sharing is related with an increase of a 27.9% likelihood of complying with medium effort offers, which is significant at a 5% level. The increase of the likelihood of compliance with high effort offers when information is shared is not significant.

Table 8 shows the relationship between the wage offered and the rejecting and compliance behaviour

Dependent variable: Compliance	(1) (all)	(2) (all)	(3) (high)	(4) (med)	(5) (med+high)
Amount offered	0.0127***	0.0108***	0.00845*	0.0165*	0.0113***
	(0.00247)	(0.00320)	(0.00392)	(0.00815)	(0.00343)
Amount offered squared					
HIGH demanded?	-0.848***	-0.924***			-0.270***
	(0.0983)	(0.0969)			(0.0651)
MEDIUM demanded?	-0.559***	-0.662***			
	(0.0941)	(0.101)			
Information sharing?	0.0810	-0.194*	-0.0537	0.226	0.0634
	(0.0472)	(0.0935)	(0.154)	(0.192)	(0.0877)
Information sharing? × Amount offered		0.00558	0.00681	-0.00578	0.00701
		(0.00487)	(0.00607)	(0.0128)	(0.00501)
Information sharing? × HIGH demanded?		0.188			-0.109
		(0.148)			(0.0967)
Information sharing? × MEDIUM demanded?		0.279**			
		(0.108)			
Constant	0.897***	1.004***	0.161*	0.237*	0.342***
	(0.0635)	(0.0674)	(0.0819)	(0.112)	(0.0586)
Observations	769	769	396	288	684
R-squared	0.555	0.563	0.657	0.606	0.537
Adjusted R-sq	0.498	0.505	0.561	0.451	0.469
Period dummies	Yes	Yes	Yes	Yes	Yes
Worker fixed effects	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The relationship between wage and compliance by the worker in treatments (3E) and (3ES). Compliance means that the worker chooses the effort level specified by the employer, or a higher effort level. Information sharing is an indicator variable indicates that the player was in the treatment with information sharing (3ES). Column (1) and (2) shows the relationship for all wage offers, while columns (3), (4) and (5) show the relationship for all offers where the employer demanded respectively high, medium and medium or high effort. Worker fixed effects are included. As in Figure 5, note that for accepted offers demanding low effort, compliance is certain, as any effort choice complies with the effort demanded by the employer. For this reason offers with a low effort demand are excluded in columns (3), (4) and (5).

	(1) Wage offered	(2) Wage offered	(3) Wage offered
Prev. compliance?	0.846	0.853	0.837
	(1.93)	(1.95)	(1.83)
Prev. contracted?	-0.435	-0.780	-0.773
	(-1.07)	(-1.18)	(-1.15)
Prev. compliance for others?		0.514	0.401
		(1.57)	(0.83)
Prev. contracted with other?		-0.629	-0.506
		(-0.90)	(-0.66)
Information sharing?	-0.187	-0.209	-0.0647
	(-0.16)	(-0.17)	(-0.06)
Information sharing? $\times$ Prev. compliance for others?			0.415
			(0.41)
Information sharing? $\times$ Prev. contracted with other?			-0.464
			(-0.71)
Constant	14.91***	15.24***	15.20***
	(34.29)	(22.77)	(22.79)
Observations	1795	1795	1795
R-squared	0.698	0.699	0.699
Adjusted R-sq	0.682	0.682	0.681
Period dummies	Yes	Yes	Yes
Employer fixed effects	Yes	Yes	Yes

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: The link between the wage offered and compliance in the previous period in treatments (3E) and (3ES). Information sharing is an indicator variable indicates that the player was in the treatment with information sharing (3ES). Stars indicate the p-values of the robust standard errors, clustered at the session-game level.

in the previous period, separated for whether the worker was working for the employer who makes the wage offer or for another employer. Again, we do not find evidence for conditional reciprocity: none of the coefficients are significant at a 5% level. Furthermore, we do not find a significant relationship between information sharing and the wage offered.

5 Discussion and conclusion

5.1 Heterogeneity of workers: a classification of types

In our results we document a significant drop in effort provision as soon as imperfect enforcement was introduced, leading to a dramatic drop in the earnings of the employers. While employers are making generous offers, the workers do not seem to reciprocate this by choosing high effort. On average, this seems to be at odds with predictions of models emphasizing fairness concerns. However, this does not rule out that there could be a significant amount of heterogeneity between workers.

In this section we combine the choices made by the workers in game 2 of treatment (1E) with the Fehr-Schmidt model of inequality aversion, to derive bounds for the fairness parameters in their model. In the Fehr-Schmidt model, there are two individual fairness parameters: α_i, which is the disutility from receiving a lower wage than the other agent (disadvantageous inequality), and β_i which is the disutility

from receiving a higher wage than the other agent (advantageous inequality). When $\alpha_i = \beta_i = 0$, the agent only cares about his or her own payoffs (which equals the "rational" framework).

To construct these bounds, we assume that a worker, when making a decision, only takes the payoffs of that particular period into account. We then calculate for each choice in each period the possible set of values of α_i and β_i for which this choice would indeed have maximized utility, using the Fehr-Schmidt preference function (see Figure 1 and 2 for examples of value sets). Next, for each worker in each five period game, we combine these possible values of α_i and β_i and test to what extent they overlap. The area of overlap gives us an idea of the bounds of α_i and β_i for a particular worker (see Appendix Section A.4 for graphical examples).

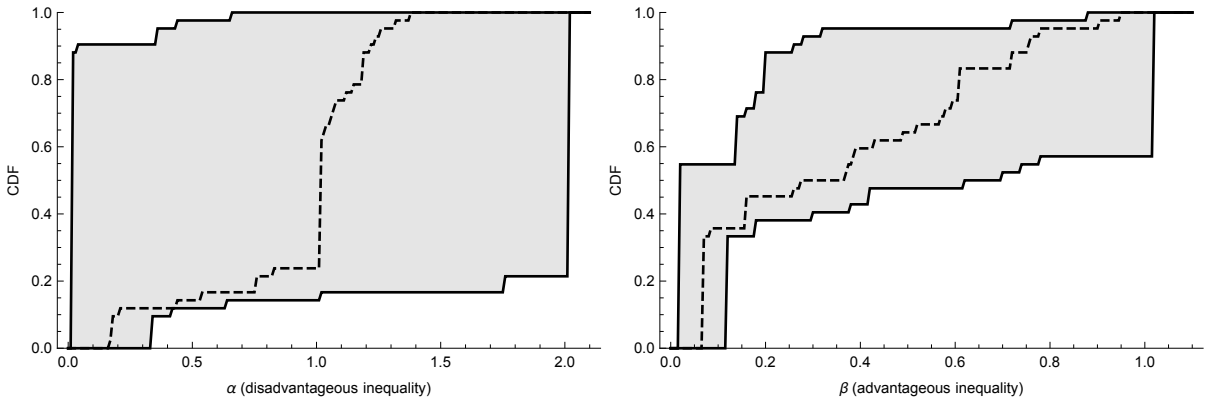
If we require the possible value sets of α_i and β_i to overlap for all five periods, we find that 33.9% of the workers satisfy this requirement. If we are a bit more lenient and allow for a "mistake", and look for values of α_i and β_i where the value sets overlap in four out of five periods, we find an overlapping value set for 71.2% of the workers.

Figure 6 shows the potential cumulative distribution functions for α_i and β_i that follows from this procedure. Only the workers for which we have found an overlapping value set in four out of five periods are included. We find that the choices of the workers are not particularly informative about the distribution of α_i, the coefficient of disadvantageous inequality aversion. The choices of workers are more informative about the distribution of β_i, the coefficient of advantageous inequality aversion.[11]

The potential cumulative distribution functions shows that there is substantial heterogeneity in the β_i values of the workers. Figure 6 shows that for 45% of the workers included the minimum value of β_i is 0.14 or higher. These workers could be typed as "fair" in the Fehr-Schmidt framework: a β_i of 0.13 corresponds to a choice of either medium or high effort (see Figures 1 and 2). For about 30% of the workers, the maximum value of β_i is 0.13 or lower. These workers could be described as more "selfish" in the Fehr-Schmidt framework. For the remaining 15% workers, the choices made by these workers do not allow for putting such a clear minimum or maximum bound on β_i. Nevertheless, this analysis shows that there is a large degree of heterogeneity: there is a substantial group of workers that can be described as "fair" and a substantial group of workers that can be described as more "selfish".

But how do employers respond to workers with low values of β_i and workers with high values of β_i? Figure 7 shows the wage offered and the earnings of the employers, grouped by whether we can classify them as having a β_i below 0.13 or a β_i above 0.13. We see that the payoffs to the employers are drastically different from each other. However, the wages offered to both groups are very similar: even though the workers with a β_i below 0.13 will mostly choose low effort, the wages are still substantially higher than the payoff that the employer gets if the worker choose low effort (five points). We find therefore little evidence of adaptation on behalf of the employers, and little punishment behaviour.

[11] This is consistent with what Charness & Haruvy (2002) conclude, that in gift-exchange games only β_i can be estimated meaningfully.



(a) Distribution of α_i **(b)** Distribution of β_i

Figure 6: The cumulative distributions of α_i, the coefficient for disadvantageous inequality aversion, and β_i, the coefficient for advantageous inequality aversion, based on the choices made by workers in periods 6-10 of treatment (1E). The gray area indicate where the cumulative distribution function could lie. The black solid lines represent the cumulative distribution of the minimum and maximum and the black dotted line the cumulative distribution of the average of the set of feasible values. Only workers for whom a minimum of four out of five choices could be rationalized with a particular set of values of α_i and β_i are included (71.2% of all workers).

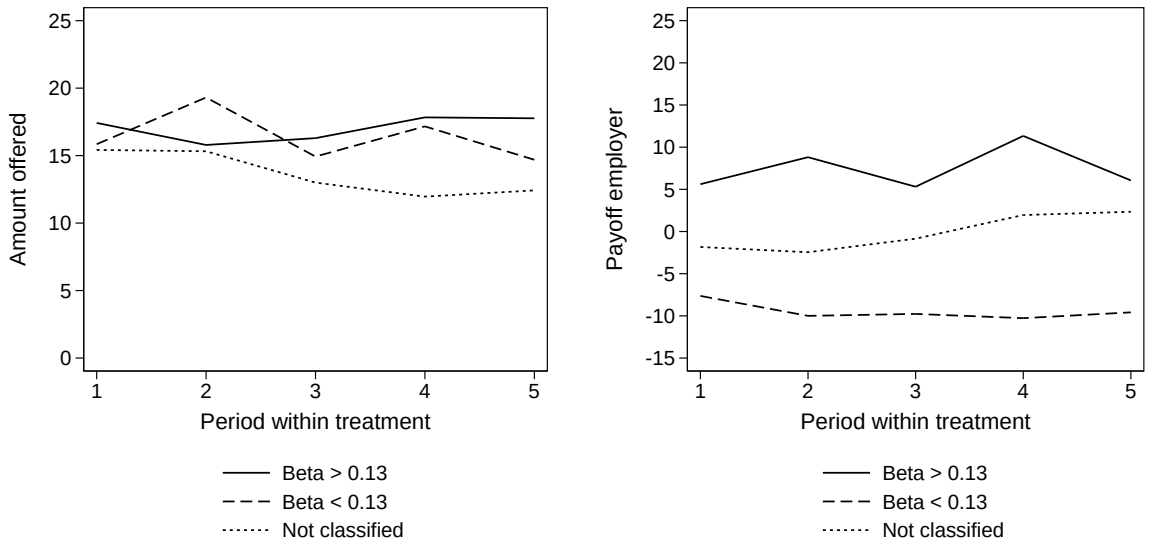


Figure 7: The wages offered and the earnings of the employers for the two classified types of workers and for the remaining workers.

5.2 Concluding remarks

Earlier experiments with gift-exchange games in developed countries have shown strong support for cooperating behaviour and relational contracting, based on reciprocity. In our experiment in Ghana we find low levels of efficiency and low levels of cooperation, especially on the side of the worker. A substantial group of workers chooses low effort, even after receiving "fair" offers, where a large share of surplus is given to the worker. We find a substantial degree of worker heterogeneity. We find little evidence for punishment behaviour on behalf of the employers: while their economic loss of a worker choosing low effort instead of high effort is 35 points, there is no significant decrease in the wage they offer to a worker choosing low effort in the next period. On average, despite low compliance, employers keep offering high wages, proposing an equal or better split of surplus.

Even though the low efficiency is in line with firm surveys done in developing countries, the lack of punishment on behalf of the employers is surprising. It is not in line with both predictions of prominent theoretical models of cooperation and gift-exchange game experiments done in developed countries: in (Brown et al., 2004) and (Brown et al., 2012) employers punish poorly performing workers, even in a situation of a shortage of labour.

We believe that our results are not the result of a lack of understanding by the participants. In the game, we tried to make the payoffs under the different effort levels as salient as possible, by explicitly showing the possible payoffs under different levels of effort while making the offer. Also, in the rehiring stage we explicitly ask the employers to make conditional offers based on the effort chosen of the worker, encouraging the employers to think about what they would offer in the next period. Despite this encouragement, we do not find strong evidence for conditional contracting in the offer stage of the game.[12] In a follow-up experiment focusing on the role of communication and feedback, we use a simplified version of the game and obtain a similar lack of punishment (Davies & Fafchamps, 2015). In these experiments, employers are more likely to send a negative feedback message after the worker has chosen low effort, indicating that employers knew that the worker was performing badly. Also, for these experiments, we invited entrepreneurs from small and medium sized enterprises to participate as well. We do not find significant differences between the behaviour of students and entrepreneurs.

The issue of framing could be important here: our games are explicitly framed in labour market terms. This was mainly done to increase understanding of the game. However, introducing such a frame comes with strong (cultural) connotations of employers providing a living to employees. Even though roles are randomly assigned, by giving somebody the identity of being an employer and somebody else the identity of a worker, these identities could affect their fairness concerns. An employer

[12]Note that in the rehiring stage, we do find some evidence for conditional reciprocity in some treatments. However, the difference between offers after low and high effort provision is small compared to the losses faced by the employer. See also appendix section A.3.

could be seen as a person who is well off, which then could reduce the reciprocity motivations of the worker.

Another related point is that employers do not see lowering a wage or terminating the relationship as a punishment. Instead, they might vent their frustration in different ways, by for example cursing at the person, or shaming the worker in a public environment. In Davies & Fafchamps (2015) we explore this further.

References

Altmann, S., Falk, A., Grunewald, A., & Huffman, D. (2014). Contractual Incompleteness, Unemployment, and Labour Market Segmentation. *Review of Economic Studies*, 81(1), 30–56.

Andreoni, J. & Miller, J. H. (1993). Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *Economic Journal*, 103(418), 570–585.

Axelrod, R. (1984). *The Evolution of Co-operation*. New York, NY: Basic Books.

Berg, J., Dickhaut, J., & McCabe, K. (1994). Trust, reciprocity and social history. *Games and Economic Behavior*, 10, 122–142.

Bloom, N., Lemos, R., Sadun, R., Scur, D., & Van Reenen, J. (2014). JEEA-FBBVA Lecture 2013: The new empirical economics of management. *Journal of the European Economic Association*, 12(August 2014), 835–876.

Bloom, N. & Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics*, 122(4), 1351–1408.

Bohnet, I., Herrmann, B., & Zeckhauser, R. J. (2010). Trust and the reference points for trustworthiness in Gulf and Western countries. *Quarterly Journal of Economics*, 125(2), 811–828.

Bolton, G. E., Katok, E., & Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. *Journal of Public Economics*, 89(8), 1457–1468.

Brandts, J. & Charness, G. (2004). Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence. *The Economic Journal*, 114(1999), 684–708.

Brown, M., Falk, A., & Fehr, E. (2004). Relational Contracts and the Nature of Market Interactions. *Econometrica*, 72(3), 747–780.

Brown, M., Falk, A., & Fehr, E. (2012). Competition and Relational Contracts: the Role of Unemployment As a Disciplinary Device. *Journal of the European Economic Association*, 10(4), 887–907.

Camerer, C. & Weigelt, K. (1988). Experimental Tests of a Sequential Equilibrium Reputation Model. *Econometrica*, 56(1), 1.

Cardenas, J. C. & Carpenter, J. (2008). Behavioural Development Economics: Lessons from Field Labs in the Developing World. *Journal of Development Studies*, 44(3), 311–338.

Charness, G., Du, N., & Yang, C.-L. (2011). Trust and trustworthiness reputations in an investment game. *Games and Economic Behavior*, 72(2), 361–375.

Charness, G. & Haruvy, E. (2002). Altruism, equity, and reciprocity in a gift-exchange experiment: an encompassing approach. *Games and Economic Behavior*, 40(2), 203–231.

Davies, E. & Fafchamps, M. (2015). Pledging, Praising, Shaming: Experimental Labour Markets in Ghana. *CSAE Working Paper*.

Fafchamps, M. & Söderbom, M. n. (2006). Wages and Labor Management. *Journal of Human Resources*, 41(2), 346–379.

Fehr, E. & Falk, A. (1999). Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy*, 107(1), 106–134.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *The Quarterly Journal of Economics*, 108(2), 437–459.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42, 1–34.

Fehr, E. & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114(3), 817–868.

Fudenberg, D. & Maskin, E. (1986). The Folk Theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3), 533–554.

Ghosh, P. & Ray, D. (1996). Cooperation in Community Interaction without Information Flows. *The Review of Economic Studies*, 63, 491–519.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *Quarterly Journal of Economics*, 115(3), 811–846.

Greif, A. (2000). The fundamental problem of exchange: A research agenda in Historical Institutional Analysis. *European Review of Economic History*, 4(3), 251–284.

Güth, W., Schmittberger, R., & Schwarze, B. (1983). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 367–388.

Güth, W. & Tietz, R. (1990). Ultimatum Bargaining Behaviour: A survey and comparison of experimental results. *Journal of Economic Psychology*, 11, 417–449.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, 33, 61–83; discussion 83–135.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–70.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining. *Journal of Economic Theory*, 104(1), 16–47.

Karlan, D. S. (2005). Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *American Economic Review*, 95(5), 1688–1699.

Kranton, R. E. (1996). Reciprocal Exchange: A Self-Sustaining System. *American Economic Review*, 86(4), 830–851.

Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational Cooperation in the Finitely Dilemma Repeated Prisoners' Dilemma. *Journal of Economic Theory*, 252(27), 245–252.

Milgrom, P. R., North, D. C., & Weingast, B. R. (1990). The Role of Institutions in the Revival of Trade: the Law Merchant, Private Judges, and the Champagne Fairs. *Economics and Politics*, 2(1), 1–23.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and Market Behavior in Jerusalem, Ljubljana , Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review*, 81(5), 1068–1095.

Selten, R. & Stoecker, R. (1986). End behavior in sequences of finite Prisoner's Dilemma supergames. A learning theory approach. *Journal of Economic Behavior & Organization*, 7(1), 47–70.

Shapiro, C. & Stiglitz, J. E. (1984). Equilibrium as a Worker Unemployment Discipline Device. *American Economic Review*, 74(3), 433–444.

A Appendix

A.1 Extensive-form representation

Figure 8 shows an extensive form representation of the game.

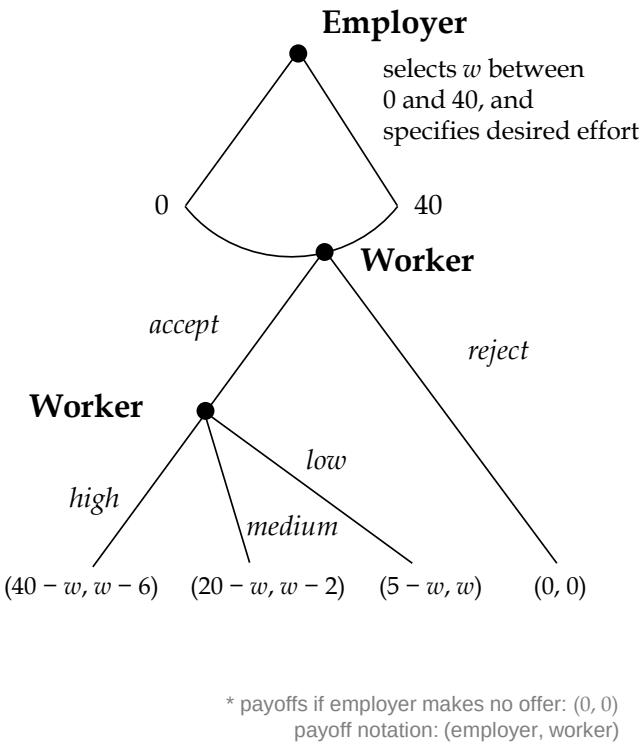


Figure 8: Extensive-form representation of the game

A.2 Reduced-form estimates

To assess the effect of the treatments, we run reduced-form regressions on two of the variables of interest: the total surplus and the wage. We run both regressions with and without worker or employer fixed effects, and cluster standard errors at the session-game level. Table 9 shows the coefficients of these regressions. The baseline treatment (1C) in game 1 is taken as the reference category of the indicator variables. Table 10 and 11 test for differences between the coefficients of the fixed effects regressions, using a t-test on the coefficients.

	(1) Surplus	(2) Surplus	(3) Wage	(4) Wage
Constant	21.32***	21.32***	20.02***	19.60***
	(0.688)	(0.632)	(0.858)	(0.650)
(1C) × Game 1	(reference)	(reference)	(reference)	(reference)
	(n/a)	(n/a)	(n/a)	(n/a)
(1C) × Game 2	3.278	1.343	-0.0599	-2.094*
	(4.104)	(4.624)	(1.526)	(1.250)
(1C) × Game 3	6.107***	4.171*	0.145	-1.843
	(1.441)	(2.365)	(1.455)	(1.279)
(1C) × Game 4	6.650***	4.714**	1.340	-0.694
	(1.240)	(2.246)	(1.185)	(1.820)
(1E) × Game 2	-8.685***	-7.785***	-4.684**	-4.294***
	(1.277)	(1.027)	(1.992)	(0.926)
(1E) × Game 3	-10.86***	-9.758***	-6.752***	-6.857***
	(0.991)	(1.195)	(2.086)	(1.120)
(1E) × Game 4	-11.35***	-10.25***	-6.895**	-6.976***
	(1.157)	(1.146)	(2.741)	(1.114)
(3C) × Game 2	4.855***	4.422***	0.384	0.943
	(1.402)	(1.036)	(1.587)	(0.815)
(3C) × Game 3	4.585**	5.938**	0.584	-0.323
	(2.250)	(2.465)	(4.369)	(1.104)
(3C) × Game 4	3.492**	4.844***	-0.247	-0.547
	(1.517)	(1.390)	(2.208)	(1.669)
(3E) × Game 3	-8.179***	-8.361***	-5.324***	-4.360***
	(1.437)	(1.156)	(1.262)	(0.777)
(3E) × Game 4	-9.027***	-9.079***	-8.275***	-6.345***
	(2.063)	(1.491)	(1.945)	(0.974)
(3ES) × Game 4	-7.813***	-8.093***	-6.769***	-6.566***
	(1.032)	(1.203)	(1.623)	(0.848)
Observations	2660	2660	4750	4750
R-squared	0.190	0.317	0.109	0.630
Adjusted R-sq	0.186	0.278	0.107	0.618
Fixed effects	None	Worker	None	Employer

Standard errors in parentheses (clustered on the session game level)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Reduced-form regressions on surplus and wage, with and without fixed effects

Game	(IC) 1	2	3	4	(IE) 2	3	4	(3C) 2	3	4	(3E) 3	4	(3ES) 4
(1C)-g1	0.0												
(1C)-g2	1.3	0.0											
(1C)-g3	4.2*	2.8	0.0										
(1C)-g4	4.7**	3.4	0.5	0.0									
(1E)-g2	-7.8***	-9.1*	-12.0***	-12.5***	0.0								
(1E)-g3	-9.8***	-11.1**	-13.9***	-14.5***	-2.0	0.0							
(1E)-g4	-10.2***	-11.6**	-14.4***	-15.0***	-2.5**	-0.5	0.0						
(3C)-g2	4.4***	3.1	0.3	-0.3	12.2***	14.2***	14.7***	0.0					
(3C)-g3	5.9**	4.6	1.8	1.2	13.7***	15.7***	16.2***	1.5	0.0				
(3C)-g4	4.8***	3.5	0.7	0.1	12.6***	14.6***	15.1***	0.4	-1.1	0.0			
(3E)-g3	-8.4***	-9.7**	-12.5***	-13.1***	-0.6	1.4	1.9	-12.8***	-14.3***	-13.2***	0.0		
(3E)-g4	-9.1***	-10.4**	-13.3***	-13.8***	-1.3	0.7	1.2	-13.5***	-15.0***	-13.9***	-0.7	0.0	
(3ES)-g4	-8.1***	-9.4*	-12.3***	-12.8***	-0.3	1.7	2.2	-12.5***	-14.0***	-12.9***	0.3	1.0	0.0

Table 10: The difference in coefficients between treatments from the fixed effects regression on surplus from Table 9. The average surplus of the (1C) treatment in game 1 is 21.3. The stars indicate the *p*-values from a *t*-test on the significance of the difference. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Game	(IC) 1	2	3	4	(IE) 2	3	4	(3C) 2	3	4	(3E) 3	4	(3ES) 4
(1C)-g1	0.0												
(1C)-g2	-2.1*	0.0											
(1C)-g3	-1.8	0.3	0.0										
(1C)-g4	-0.7	1.4	1.1	0.0									
(1E)-g2	-4.3***	-2.2	-2.5	-3.6*	0.0								
(1E)-g3	-6.9***	-4.8***	-5.0***	-6.2***	-2.6**	0.0							
(1E)-g4	-7.0***	-4.9***	-5.1***	-6.3***	-2.7**	-0.1	0.0						
(3C)-g2	0.9	3.0**	2.8*	1.6	5.2***	7.8***	7.9***	0.0					
(3C)-g3	-0.3	1.8	1.5	0.4	4.0***	6.5***	6.7***	-1.3	0.0				
(3C)-g4	-0.5	1.5	1.3	0.1	3.7**	6.3***	6.4***	-1.5	-0.2	0.0			
(3E)-g3	-4.4***	-2.3	-2.5*	-3.7*	-0.1	2.5**	2.6**	-5.3***	-4.0***	-3.8**	0.0		
(3E)-g4	-6.3***	-4.3***	-4.5***	-5.7***	-2.1**	0.5	0.6	-7.3***	-6.0***	-5.8***	-2.0***	0.0	
(3ES)-g4	-6.6***	-4.5***	-4.7***	-5.9***	-2.3***	0.3	0.4	-7.5***	-6.2***	-6.0***	-2.2***	-0.2	0.0

Table 11: The difference in coefficients between treatments from the fixed effects regression on the wage offered from Table 9. The average wage in the (1C) treatment in game 1 is 19.6. The stars indicate the *p*-values from a *t*-test on the significance of the difference. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.3 Rehiring behaviour of employers

Dependent variable: Wage offered	(1) (1E)	(2) (3E)	(3) (3ES)
Medium effort exerted	4.139**	0.0258	0.650
	(1.230)	(0.343)	(1.521)
High effort exerted	6.749*	0.173	-0.640
	(2.855)	(0.320)	(1.305)
Constant	13.41***	13.38***	12.95***
	(1.750)	(0.169)	(0.626)
Observations	133	579	309

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: The amount offered by the employer to the worker in the rehiring stage in treatments (1E), (3E) and (3ES). Standard errors are clustered on the session level.

In the rehiring stage, we ask the employers for each of the possible effort levels whether they would like to make an offer to the worker again, and what the offer would be. Similarly, the workers are asked

for their minimum wage in order to accept an offer. On average, 27.4% of the employers specify offers in the rehiring stage. Only a small share of trades is initiated through this rehiring: around 2-3 percent.

The Table above shows a regression of the amount offered by the employer to the worker, with the level of effort chosen by the worker as one of the regressors. In the (1E) treatment, we find that high effort is rewarded with a wage that is 6.7 points higher than a low effort offer (significant at the 10% level). A medium effort offer is rewarded with a wage that is 4.2 points higher and is significant at the 5% level. We do not find significant differences in the (3E) and (3ES) treatments.

These results show that there is some presence of conditional contacting by employers. However, the differences in wages are small and not sufficient to cover the losses incurred by low effort provision (the difference in the payoff to the employer between high and low effort provision is 35 points).

A.4 Classifying worker types

We assume that in every period the workers choose according to Fehr-Schmidt preferences:

$$U_i(x) = x_i - \alpha_i \max(x_j - x_i, 0) - \beta_i \max(x_i - x_j, 0)$$

Figure 1 and 2 show the areas of α_i and β_i that are consistent with making a particular choice. For each worker, using these sets of values, we can take his or her choices, and plot the value sets of α_i and β_i that are consistent with this choice for the five periods in each game:

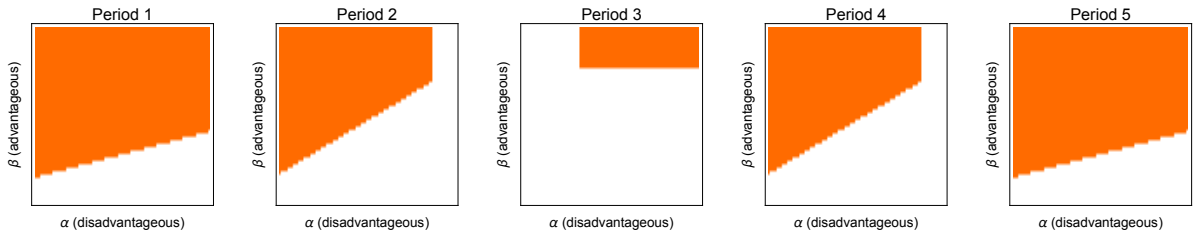


Figure 9: The value sets of α_i and β_i consistent with the choice made by one worker in the five periods of a game.

Next, we can overlap the areas with each other and then consider the set of values which are consistent with at least four out of the five choices made by a worker:

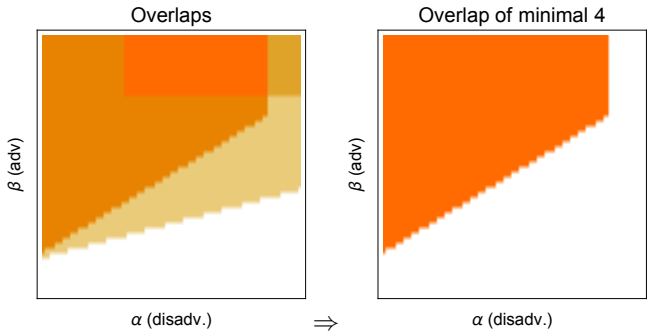


Figure 10: The overlapping area of α_i and β_i for one worker.

We repeat this process for each worker, and using this way, derive bounds for α_i and β_i.