Deferred Compensation in Multi-Period Labor Contracts: An Experimental Test of Lazear's Model

Steffen Huck, University College London, ELSE, IZA
Andrew J. Seltzer, Royal Holloway, University of London
Brian Wallace, University College London

We gratefully acknowledge a grant from the Nuffield Foundation. In addition, the first author acknowledges financial support from the Economic and Social Research Council (UK) via the Centre for Economic Learning and Social Evolution (ELSE). We have benefited from comments by Dirk Engelmann, Michael Naef, Hans Normann and participants in seminars at Royal Holloway, Colby College, UCL, IZA, and the Society of Labor Economists Conference.

Deferred Compensation in Multi-Period Labor Contracts: An Experimental Test of Lazear's Model

Abstract

This paper provides the first experimental test of Edward Lazear's (1979) model of deferred compensation. We examine the relationship between firms' wage offers and workers' effort supply in a multi-period environment. If firms can *ex ante* commit to a wage schedule with deferred compensation, workers should respond by supplying sufficient effort to avoid dismissal. We contrast this full-commitment case to controls with no commitment and computer-generated wages in order to examine the roles of monetary incentives, social preferences, and reciprocity. Finally, we examine a setup without formal commitment, but where firms can build a reputation for paying deferred wages.

Keywords: deferred compensation, experimental labor economics, personnel economics, social preferences, incentives.

JEL class: C91, J31, J41, M51, M52.

Few questions in labor economics have been the subject of as much attention as the effect of workplace incentives on worker effort (Edward Lazear, 2000 and Canice Prendergast, 1999). While this literature has traditionally combined theory with the empirical analysis of observational data, the last decade has seen an ever-increasing number of experimental studies shedding light on the nexus between incentives and effort. Important contributions include Ernst Fehr, Georg Kirchsteiger, and Arno Riedl (1993) on the role of gift exchange and Fehr, Simon Gachter, and Kirchsteiger (1997) on the role of reciprocity. More recently, this literature has focussed on the issue of dynamic, long-term 'relational' contracts. Specifically Martin Brown, Armin Falk, and Fehr (2004 and 2008) study repeated labor markets where agents are identifiable and workers and firms can endogenously match into long-term pairs. Such endogenous relations are shown to be substantially efficiency enhancing. However, while worker-firm relations are dynamic in these studies, contracts as such are always for one period only.

In this study we examine multi-period labor contracts in order to study the role of deferred compensation, as described by Lazear (1979). Deferred compensation is likely to emerge if the worker-firm relationship is expected to be long-term and direct monitoring of worker productivity in the short-term is either costly or not possible. In a deferred compensation contract, workers are underpaid during the early part of their career (that is, pay is less than marginal revenue product) and overpaid during the later part of their career. This structure of compensation encourages higher effort because future pay within the firm always exceeds future pay elsewhere.[1] An important issue concerning deferred compensation is that firms have an incentive to renege on late-career overpayment, as older workers are paid more than their marginal product. This implies that in the absence of an effective commitment mechanism, deferred compensation cannot be part of an equilibrium contract. The literature

[1] Deferred compensation contracts have also been shown to reduce quits and attract future-minded employees (Steven Salop and Joanne Salop, 1976). We do not analyze these properties in this paper.

has focused on two such mechanisms: legally binding long-term contracts and reputation (firms that dismiss older workers will face future recruiting difficulties).[2]

Because deferred compensation is an optimal contract only if there exist difficulties in measuring effort and output, there is an inherent problem in testing the theory directly on observational data. If the short-term performance were directly observable then the employer should be able to optimally tie wages to productivity in the short-run through, for example, piece rates. So, unsurprisingly, we are not aware of any direct empirical evidence on the impact of deferred compensation on worker behavior. Instead the existing literature focuses on whether mandatory retirement and steep age earnings profiles are associated with jobs with monitoring difficulties (Robert Hutchens, 1987); whether mergers or high bankruptcy probabilities are associated with flatter wage profiles (Jagadeesh Gokhale et. al., 1995); and whether wages increase too much in later years of the career to be explained by productivity increases (Seltzer and David Merrett, 2000).

This paper provides the first direct experimental test of Lazear's model. We create a simple three-period model in which positive worker effort is a socially optimal outcome. Effort levels cannot be contracted on directly; however, low effort by the worker can result in dismissal with some probability. The model captures several of the important equilibrium features of the Lazear model. Assuming that firms are able to commit to wage offers in advance, young workers will earn less than their marginal product, older workers will earn more than their marginal product, and firms will offer deferred compensation in order to increase worker effort. This is the scenario we investigate in our main experimental treatment. In addition, we study two control treatments. The first additional treatment controls for the importance of binding monetary incentives, by removing the firm's commitment power. In this scenario, orthodox theory predicts a complete breakdown of the worker-firm relationship with zero wages and zero effort. The second additional treatment controls for social

[2] A second issue is that older workers will have an incentive to remain on the job past the optimal retirement point because they are paid more than marginal revenue product. This problem can be solved by a policy of mandatory retirement.

preferences and reciprocity in workers' behavior by replacing wage offers from human firm subjects with computer-generated wage offers. While the first three treatments are designed to provide an ideal model *test*, the final treatment that we introduce is of a more explorative nature. In this treatment also, firms do not have formal commitment power but can build up a reputation for paying deferred wages as their past histories of play are made available to workers. This treatment allows us to gain some insight whether under more realistic conditions deferred compensation can emerge endogenously.

The remainder of the paper is organized as follows. The first section outlines a simple three-period model that captures the essential features of Lazear's deferred compensation model. The second section describes the design and setup of our experiment. The third section uses the model to make predictions for the experiment. In addition to considering the orthodox case of pure income-maximization, we also outline predictions for the model under the assumption that agents have social preferences regarding inequity aversion, in the manner described by Fehr and Klaus Schmidt (1999). The fourth section presents the results from the experiment, and analyzes the results for both worker and firm players in the context of the model. The fifth section concludes.

I. The Model

Figure 1 shows a deferred compensation contract. In a competitive labor market, the contract offers lifetime wages that are equal in net present value to lifetime productivity; however, wages are less than productivity prior to T* and greater than productivity thereafter.[3] These contracts need two additional features to provide optimal incentives. In order for lifetime wages to be no greater than lifetime productivity, the worker must retire at or before T**. At the start of the contract, the worker is indifferent to the wage payments shown with retirement at T** and a contract with wage equal to MRP. However, since the wage is greater than the worker's value of leisure at T**, he will continue working unless the

[3] The competitive labor market assumption is not necessary to capture the main points of the Lazear model. Our model does not assume competitive labor markets, and in equilibrium lifetime wages are less than lifetime productivity.

contract contains a mandatory retirement provision. It is also necessary to have provisions that make dismissal of older workers difficult in order to prevent firms from reneging on the implicit contract and dismissing the worker at T*. Lazear offers two mechanisms for this purpose: legally binding contracts such as "last hired, first fired" and informal means that depend on maintaining an honest reputation in order to recruit workers in the future.

Our experiments use a discretized version of Lazear's continuous time model with three periods reflecting the phases before T* (young worker – stage 1), between T* and T** (old worker – stage 2), and after T** (retired worker – stage 3).

At the start of each round, firm players have to decide about wages for the first two stages and about a pension for the third:

$$(W_1, W_2, W_3) \in R_+^3$$

In our main treatment, the Full Commitment Treatment (FCT), we assume that these offers are binding for all three stages. Worker players then have to decide their effort level in stages 1 and 2, from Low, Medium or High:

$$E_1, E_2 \in \{L, M, H\}$$

These three effort levels represent effort so low that workers becomes sackable, effort that can be financially incentivized, and non-incentivizable effort that can only be reached if there is some of form "gift exchange" as documented in worker-firm relations by Fehr, Kirchsteiger, and Riedl (1996).

At each stage, the worker's effort translates to output, Z_{Ei} for the firm with:

$$Z_H > Z_M > Z_L$$

We assume that there is an increasing cost of effort, C_{Ei}, i.e.

$$C_H > C_M > C_L$$

This holds in both stages. If (and only if) low effort is detected, the worker is either sacked (after stage 1) or loses his pension (after stage 2), although he is still paid his earnings for the current stage. For the sake of simplicity, the detection probability is exogenously determined to be *p*. Think of this as the optimal detection technology available to the firm. We assume that the firm can make a full commitment not to renege and dismiss a worker who has supplied medium or high effort.

Formally, the structure of the game and payoffs to the firm and worker are as follows.

Stage 1: The firm chooses a wage offer (W_1, W_2, W_3). After the worker sees this, he chooses $E_1 \in \{L, M, H\}$. If $E_1 = L$, then with probability *p* the worker is sacked and the game ends. In this case, the firm payoff is $2Z_L - W_1$ (the firm's output in both stages is Low) and the worker's payoff is $W_1 - C_L$.[4]

If the worker is not sacked at stage 1, the game proceeds to stage 2.

Stage 2: The worker chooses $E_2 \in \{L, M, H\}$. If $E_2 = L$, then again with probability *p* the worker is sacked and the game ends. In this case, the firm's payoff is $Z_{E1} + Z_L - W_1 - W_2$ and the worker's payoff is $W_1 + W_2 - C_{E1} - C_L$

If the worker is not sacked at stages 1 or 2, the game proceeds to stage 3.

Stage 3: No choices are made by firm or worker. The pension, W_3, is paid. The firm payoff is $Z_{E1} + Z_{E2} - W_1 - W_2 - W_3$ and the worker's payoff is $W_1 + W_2 + W_3 - C_{E1} - C_{E2}$

[4] We assume that if the worker is sacked, the firm's output in stage 2 is the same as if the worker had supplied low effort.

In addition, we add to the worker pay a constant, I, representing the outside lifetime income, regardless of when the game ends.

As an initial benchmark, we assume that workers are pure income-maximizers. We later revisit this assumption and consider the role of social preferences. The equilibrium is as follows. First, note that in a one-shot game it is off the equilibrium path for the worker ever to supply high effort, even if that is the socially optimal outcome, because the cost to the worker is increasing in effort level and he cannot be dismissed for supplying medium effort.[5] The subgame perfect equilibrium solution is found by observing that if:

$$(Z_M - C_M) > (Z_L - C_L)$$

the joint surplus is higher if the worker supplies medium effort in both periods than if he supplies low effort. Thus the firm will be willing to compensate the worker for the cost of extra effort. Analyzing the worker's second-period choice of whether to supply medium effort, we find he does so whenever:

$$(1) \qquad (1 - p)W_3 - C_L \leq W_3 - C_M$$

His first-period choice depends on what he expects to do in the second period. If condition (1) is fulfilled, i.e. if the pension is sufficiently high to deter low effort in the second period, then medium effort in stage 1 pays only if:

$$(2) \qquad (1 - p)\,(W_2 + W_3 - C_M) - C_L \leq W_2 + W_3 - 2\,C_M$$

Profit maximization implies that W_2 and W_3 are minimized subject to (1) and (2). We shall refer to these conditions later summarily as the "Lazear condition". Notice that there is no restriction on W_1, which does not influence the worker's behavior and is thus zero in equilibrium.

[5] This is consistent with Lazear (2000), who argues that firms need to use additional mechanisms (such as promotion tournaments) to motivate workers to supply effort above a threshold level.

Thus the equilibrium wages can be characterized as (0, X, Y) where $0 \leq X \leq C_M$ and $Y = (C_M - C_L)/p + (C_M - X)$ and workers' equilibrium effort is (M, M). Net productivity will be $(Z_M - C_M, Z_M - C_M, 0)$ and the resulting payoffs over the three stages are $(-C_M, X - C_M, Y)$ for the worker and $(Z_M, Z_M - X , - Y)$ for the firm.

Now consider the case where the firm cannot commit in advance to its wage offer and only decides on the wage at the start of each stage. Solving by backward induction, in stage 3 the firm always chooses $W_3 = 0$, as the worker no longer produces any output. The worker anticipates this and therefore chooses to supply low effort in stage 2. The firm anticipates this and thus understands that they cannot motivate effort in stage 2. Therefore they offer the worker $W_2 = 0$, and similarly for stage 1. The equilibrium prediction is that the firm will offer the wage profile (0, 0, 0); workers will supply effort levels (L, L), and net productivity will be $(Z_L - C_L, Z_L - pC_L, 0)$. The resulting payoffs over the three stages are $(-C_L, -C_L, 0)$ for the worker and $(Z_L, pZ_L, 0)$ for the firm. Because the firm cannot commit to its offers, the end result is the Pareto inferior low-wage, low-effort equilibrium.

II. Experimental Design and Procedures

We ran four treatments in our experiment to distinguish between optimal responses to incentives, genuine mistakes, and consequences of social preferences. Our main treatment implements the Lazear case where firms can fully commit to future wages and deferred compensation arises in the subgame perfect equilibrium. We call this the Full Commitment Treatment (FCT). Our design has three further treatments. Two serve as controls for the FCT and one is of a more exploratory nature. The two controls are the No Commitment Treatment (NCT) and the Computer Firm Treatment (CFT). In the NCT, firms pay a first-stage wage but can only make non-binding promises about future payments. In the CFT, human firm subjects are replaced by computerized firms that make the same offers that we observed in the main treatment. For these three treatments we can derive clear theoretical predictions. In addition, we have a fourth exploratory treatment, the Reputation Treatment (RT) for which theoretical predictions are less clear. The RT is similar to the NCT, except that workers when matched

with a firm have access to the firm's entire history (wage promises, wages paid, and previous workers' responses), which allows for reputation building.

For the three interactive treatments (the FCT, the NCT, and the RT), ten subjects (5 workers and 5 firms) participated in each session. Subjects were assigned the role of firm or worker at the beginning of the experiment and kept their role throughout the session. Each session consisted of 20 rounds, with random matching of workers and firms which, in the case of the FCT and the NCT, approximates the nature of one-shot games in each round. The experiment was programmed and conducted with the z-Tree software (Urs Fischbacher, 2007). We ran 6 sessions for each of the treatments, and thus there are observations for 30 worker players and 30 firm players over 20 rounds for each treatment. In the fourth treatment, the CFT, there is full commitment, but the human firm players are replaced by a computer such that the CFT becomes a single-person decision problem. The 20-round wage profiles received by the worker players in the CFT were drawn from the profiles of wages received by worker players in the FCT. [6] This treatment also has 30 players and 20 rounds.[7] The subjects in each treatment were students of any field at Royal Holloway and UCL. Sessions lasted approximately 100 minutes.

In the RT, there are also 6 sessions lasting 20 rounds with 5 worker and 5 firm players. Workers and firms are also randomly matched in each round. However, firms are long-lived and carry labels, which makes them identifiable to workers. Moreover, worker players have access to firm players' entire history during the experiment. Specifically, the worker player observes the firm player's wage promises and actual payments in all previous rounds (as well as the previous worker players' responses to these). Thus, at the time they are deciding on their effort level, the worker can see if the firm has a history of making generous 2nd and 3rd stage wage offers and whether have they previously reneged on promised wages.

[6] Due to a coding error the wage profiles were not exactly the same in 21 of 600 cases for the two treatments. In most of these cases the error was in W_1. The difference in the average offer per round is only 0.2 percent (76.47 in the FCT versus 76.29 in the CFT). This does not affect any parts of our analysis.

[7] In the CFT there is no interaction between subjects and thus no need to hold constant the number of subjects in each session. We ran three sessions for this treatment, one with 6 subjects and 2 with 12 each.

The parameters of the model were set as follows: $p = 0.5$, $C_L = 0$, $C_M = 20$, $C_H = 40$, $I = 40$, $Z_L = 50$, $Z_M = 100$, $Z_H = 140$. Wage offers could range between 0 and 120 for each stage of the round.

Instructions to subjects and a sample computer screen from the RT detailing a firm player's history are available from the authors as an additional appendix.

III. Predictions Accounting for Social Preferences

In this section, we will make a number of predictions, for both the orthodox model and for scenarios with social preferences.

Given our parameter choices the worker players' Lazear optimal strategies in the FCT are:

Stage 1: E_1 = MEDIUM if $(2W_2 + W_3) \geq 80$ or $(W_2 + W_3) \geq 60$;
E_1 = LOW otherwise
Stage 2: E_2 = MEDIUM if $W_3 \geq 40$;
E_2 = LOW otherwise

The firm's equilibrium wage offer is characterized by $W_1 = 0$, $0 \leq W_2 \leq 20$, and $W_3 = (60 - W_2)$. The same observations on optimal worker behavior hold for the CFT. In contrast, in the NCT, the worker players' optimal strategies induce low effort in both stages independent of wage offers and history. The firm's equilibrium wage offer is (0, 0, 0). In the RT, the predictions of orthodox game theory depend on the modeling approach. If all agents are income-maximizers and this is common knowledge, subgame perfect equilibrium behavior is the same as in the NCT, with zero wages and low effort in both stages and all rounds. If one is willing to assume the potential existence of behavioral types for the firm player, more efficient and complex reputation equilibria, in the spirit of David Kreps et al. (1982), may arise.

Some twenty-five years after the first ultimatum game experiment (Werner Güth, et. al., 1982), it would perhaps be naïve to rely entirely on the orthodox predictions above. Our game is characterized by distributional conflicts and, by now, it is well-known that, in the presence of such conflicts, social preferences may affect behavior. Indeed, we have designed our experiment in order to be able to capture and measure such effects and contrast them to the effects of monetary incentives and pure error in decision making.

Specifically, the comparison between the FCT and NCT allows us to test for the importance of binding monetary incentives that stem from the firms' ability to commit *ex ante* to wage offers. The model presented in the previous section implies that deferred compensation and medium effort will emerge in the FCT, but not in the NCT. Accordingly, differences that arise in the two treatments can be attributed to Lazear's theory.

Similarly, the comparison between the FCT and CFT can be used to test for the importance of social preferences. As there are no other subjects making decisions or earning money in the CFT, social preferences cannot affect behavior in this treatment. Accordingly, observed differences between the FCT and CFT (that have identical wage offers) can be attributed to workers' social preferences. Moreover, the CFT informs us about the role of pure errors. Any deviation from income-maximizing behavior in the CFT must stem from some form of boundedly rational behavior or computational errors.

What then are the implications of social preferences for workers and firms? We follow a two-pronged approach to derive predictions, distinguishing between pure distributional concerns and reciprocity. To capture distributional concerns, we follow Fehr and Schmidt's (1999) model of inequity aversion. Fehr and Schmidt allow for two types of inequity aversion, aversion against disadvantageous inequity (where an agent dislikes having less than another) and aversion against advantageous inequity (where an agent dislikes having more than another). Due to the asymmetry in payoffs in our experimental game it is mainly *disadvantageous inequity aversion on the side of workers* that is expected to matter.

In order to capture (positive) reciprocity in the form of gift exchange (Marcel Mauss, 1954 and George Akerlof, 1982) we follow a non-parametric approach. Specifically, we will view a wage offer that exceeds pure incentivization and is met by higher than incentivized effort as possibly resulting from gift exchange. We will make the meaning of "exceeding pure incentivization" precise below.

First, let us consider the implications of the Fehr-Schmidt model. Formally, if we let π_F and π_W denote the firm and worker monetary payoff for the game, the Fehr-Schmidt utility for a worker is given by:

$$U_W(\pi_W, \pi_F) = \pi_W - \alpha * \max(\pi_F - \pi_W, 0) - \beta * \max(0, \pi_W - \pi_F)$$

where $\alpha \geq 0$, $\beta \leq \alpha$, and $0 \leq \beta \leq 1$ (and similarly for the firm). The standard Lazear case is embedded as the special case where $\alpha = \beta = 0$.

Notice that in the Lazear benchmark, workers earn substantially less than firms so that it is the workers' alphas and firms' betas would be expected to matter.[8] Let us focus on the workers' alphas first.

Essentially, a worker subject who dislikes earning less than the firm subject he is matched with might be tempted to "punish" the firm by exerting lower effort than an income-maximizing worker would do. Such punishments can, hence, only matter when the Lazear condition for incentivizing medium effort is met. If the income-maximizing worker would choose low effort there is no scope for further punishment, regardless of the value of alpha. Accordingly, we need to inspect wage offers that incentivize the income-maximizing worker.

Our aim is not to characterize Fehr-Schmidt equilibria for all possible distributions of alphas and betas. Rather we want to understand which deviations from the Lazear prediction might

[8] In equilibrium the firm pays $W_1 = 0$ and $W_2 + W_3 = 60$. Effort is medium in both stages leading to output of 100. Effort cost is 20 in each stage and the worker's outside earnings are 40. Thus the worker's payoff is 40 + 60 - 40 = 60. The firm's payoff is 100 + 100 - 60 = 140.

be *Fehr-Schmidt rationalizable*. In the first instance, this means simply to understand for which wages inequity averse agents *might* choose low effort where income-maximizing agents would choose medium effort. It turns out that this wage region is easy to compute. Simply notice that if the effort profile (M, M) is "Lazear incentivized" (i.e. where $W_2 + W_3 \geq 60$ and $W_3 \geq 40$) workers and firms have equal payoffs if $W_1 + W_2 + W_3 = 100$. Now consider wage offers with a total wage below that boundary such that the worker has a smaller payoff than the firm. Then the *most extreme* Fehr-Schmidt type (with an alpha approaching infinity) would "reject" the wage offer and choose low effort instead. Hence, low effort choices for total wage offers below that bound are Fehr-Schmidt rationalizable.[9]

This analysis allows us to draw a simple diagram where we divide the wage-offer space into three regions separated by two lines (see Figure 2). First, we have the *Lazear line* which runs parallel to the y-axis at x = 60, the total deferred wage that is necessary to incentivize income-maximizing workers to exert medium effort.[10] To the left of this line, the Lazear condition is not met and income-maximizing workers will choose low effort. To the right the Lazear condition is met and income-maximization prescribes medium effort. Second, we plot an iso-wage curve for a total wage of 100 to the right of the Lazear line. We shall call this the *alpha border*. To the southwest of this line, workers experience disadvantageous inequity such that their Fehr-Schmidt alphas are relevant. To the northeast of this line they experience advantageous inequity such that their alphas become irrelevant.

With the aid of Figure 2 we can now review the implications of economic and social incentives for our experimental game. Let us first make predictions for workers. In the western region (to the left of the Lazear line) all workers (regardless of their Fehr-Schmidt alphas) are not sufficiently incentivized to exert more than low effort. Medium or high effort

[9] Similarly, if (M, L) is Lazear incentivized (i.e. where $W_2 + .5W_3 \geq 40$ and $W_3 < 40$), workers' and firms' payoffs are equalized if $W_1 + W_2 + .5W_3 = 65$ and wage offers below that bound are Fehr-Schmidt rationalizable.

[10] Strictly speaking the condition for medium effort being fully incentivized is $W_2 + W_3 \geq 60$ *and* $W_3 \geq 40$. Figure 2 shows offers for which $W_1 + W_2 \geq 60$. There exist wage offers that incentive medium effort in either stage 1 or stage 2, but not both. For ease of exposition, we do not consider these offers in this exposition, but do return to them later in the analysis.

can only result from error or reciprocity and the comparison of the FCT and the CFT treatment will help us to distinguish between the two. Specifically, we expect that in this western region reciprocity might be triggered through high first-stage wages. Any wage offer in the western region only incentivizes low effort, hence any wage profile that exceeds (0, 0, 0) in this region can be viewed as a gift or an error. However, in the context of incentives provided through deferred compensation, a high first-stage wage is more clearly discernible as a gift than a high second- or third-stage wage which might have been intended for incentivization but, say, due to computational error, been chosen at too low a level. So, while both the Lazear and Fehr-Schmidt models predict low effort in the entire western region, reciprocity would predict that instances of medium effort become more prevalent the further north we move in this region.[11]

In the southern triangular region to which we shall refer as the *Fehr-Schmidt triangle*, income-maximizing workers will choose medium effort but workers with high alphas (who suffer from disadvantageous inequity) might "punish" firms and choose low effort. In other words, low effort is *Fehr-Schmidt rationalizable* in this region. The Fehr-Schmidt triangle also stands out in that it is the only wage region where fully-incentivized medium effort by the worker enables the firm to earn more than with a wage offer of (0, 0, 0) and low effort by the worker.[12]

The Fehr-Schmidt rationalizability of low offers in the triangle is in sharp contrast to the final northeastern region beyond the Lazear line and the alpha border. Here the prediction is that low effort should never occur (regardless of Fehr-Schmidt parameters). However, reciprocity might trigger high effort. The logic is similar but subtly different to the reciprocity logic discussed earlier. While offers in the Fehr-Schmidt triangle exceed the Lazear incentives, they do not exceed incentives that are necessary to induce inequity-averse agents

[11] Notice that very high wage offers ($W_1 > 65 - 0.5W_2 - 0.25W_3$) in the western region might also trigger workers who dislike advantageous inequity to balance payoffs through exerting medium or high effort.

[12] To see this, consider first the offer (0, 0, 0) and low effort in both stages. The firm earns 100. With medium effort in both stages, the firm earns $200 - (W_1 + W_2 + W_3)$. Equalizing these gives $(W_1 + W_2 + W_3) = 100$, which is the alpha border.

to exert medium effort. Thus, only offers that are beyond the alpha border *unambiguously* exceed incentivization and we expect that only then reciprocity might kick in.[13] Specifically, we might expect that workers view these offers as particularly "kind", in the sense described by Falk and Fischbacher (2006), and be more likely to reciprocate by supplying high effort.

Finally, regarding firm behavior, the predictions of the Lazear and Fehr-Schmidt models are as follows. Stage 1 wage offers have no incentive effects in either model and are, hence, set to zero in equilibrium. Thus all wage offers should be on the x-axis of Figure 2 – either on the origin where wages are (0, 0, 0) or on the bottom of the Fehr-Schmidt triangle, that is, with deferred wages between 60 and 100. The precise location would obviously depend on firms' beliefs about the distribution of alphas among workers.[14] All other wage offers contain some element of gift giving with wage offers in the northeast of the alpha border standing out as particularly generous, giving the firm a lower payout than (0, 0, 0) with medium effort.

IV. Results

A. Overview

The data from our experiments do not lend themselves to easy summary and the analysis requires some patience. We will proceed in three steps. First we will show figures plotting the *raw* data for the first stage. We will discuss these figures at some length. Then we turn to workers and firms separately, first showing some descriptive statistics, then presenting some econometric results.

[13] Again, for very high total wages, i.e., very far to the northeast, workers with sufficiently high betas might also "repay" their generous wages with higher than incentivized effort. Specifically, total wages would have to be above 130, which can be thought of as a *beta border*.

[14] In principle, one could imagine firms making very high offers that would be met by high effort from workers with sufficiently high betas. It is possible to construct Fehr-Schmidt equilibria with wages above 130 and high effort - but the equilibria require assuming distributions of alphas and betas that are extremely different from previous studies.

Figures 3A to 3D show all the key data from the first two decisions we observe in our experiments, the firm's wage offer and the worker's first-stage effort choice. The figures are constructed just like Figure 2 in our section on predictions. That is, they plot wage offers with W_1 on the y-axis and the total deferred wage, $W_2 + W_3$, on the x-axis. This wage space is divided into three regions by the *Lazear line* and the *alpha border*.

For each wage offer, the figures then plot how workers respond – with circles marking low-effort choices, triangles marking medium-effort choices and crosses marking high-effort choices. The sizes of these circles, triangles, and crosses indicate the frequency of these observations. Finally, notice that for the NCT and the RT (the latter of which we will only discuss later) the x-axis plots *promised* deferred pay rather than actual deferred pay, as this was the information available to the worker players at the time they chose their stage 1 effort level. We have retained the *Lazear line* and the *alpha border* in these diagrams as a benchmark.

Let us now examine Figures 3A – 3C, for the main treatment, the FCT, and the two controls, the CFT and the NCT. Let us begin with the FCT.

There are a number of interesting patterns that are discernible with the naked eye (and some concentration). (i) A large fraction of wage offers are, as predicted by the model, on the x-axis with the bulk on either (0, 0, 0) or in the region with deferred wages between 60 and 100. (ii) Almost all of the (0, 0, 0) wage offers generate optimal low effort. (Compare the *relative* sizes of the (0, 0, 0)-triangle and the (0, 0, 0)-circle.) (iii) In general, low effort is predominant if the Lazear condition is not met. However, for deferred pay below 60, medium effort becomes more frequent towards the north, i.e., where W_1 gets higher, suggesting gift exchange. Notice that there are also some instances of medium effort just left of the Lazear line. Some of these instances reflect a income-maximizing response to a wage offer that incentivized medium effort for the first stage only.[15] In other cases, it is likely that the subjects understood the importance of deferred compensation, but not the precise location of the Lazear line. (iv) When the Lazear condition is met, medium effort becomes much more

[15] In other words the offer was such that $W_2 + .5W_3 \geq 40$, but $W_2 + W_3 \leq 60$

frequent, and is clearly modal. However, there is a remarkable difference between the Fehr-Schmidt triangle and the northeastern region beyond the alpha border. While there are quite a few instances of low effort in the Fehr-Schmidt triangle, low effort almost disappears beyond the alpha border where workers earn at least as much as firms when they exert medium effort. This strongly suggests the importance of disadvantageous inequity aversion. (v) High effort is very rare but clearly most frequent to the northeast of the alpha border. Only if purely distributional concerns are no longer relevant, that is, if wages also exceed Fehr-Schmidt incentivization, is there scope for fully-efficient gift exchange.

Figure 3B shows our first control, the CFT, where there are only worker subjects and wage offers are made by the computer. The distribution of wage offers is, by design, identical to that of the main treatment. However, the chosen efforts are markedly different, implying a strong role for social preferences. Remember, for the CFT subjects, social preferences are irrelevant and all that should matter is their own earnings. In other words, in the CFT only the Lazear line should make a difference, the alpha border should be irrelevant and workers should *never* choose high effort. Inspecting the figure we make the following observations. (i) In general, the Lazear line organizes the data well with circles being predominant to its west and triangles predominant to its east. (ii) There is, however, some noise. There are, in particular, some instances of high effort, typically above the alpha border. This adds a slight caveat to our observations about reciprocity in the FCT. (iii) A comparison of the Fehr-Schmidt triangles in the FCT and the CFT underlines the role of disadvantageous inequity aversion. Low effort choices are, as expected, far less prevalent in the CFT where workers play a single-person game.

Let us now compare the FCT with the second control treatment, the NCT, to understand the impact of incentives and commitment. In the NCT, firms cannot commit to wages such that orthodox theory predicts firms choosing (0, 0, 0) and workers exerting low effort. Similarly, the Fehr-Schmidt model predicts low effort, although firms with sufficiently high betas may pay low, but positive wages. The dramatically different prediction is met by a dramatically different figure. The most obvious feature of the NCT diagram is the dominance of circles. Low effort is the norm regardless of W_1 and the promises made about W_2 and W_3.

There is, however, some gift exchange with medium-effort triangles appearing further towards the north. It is interesting though to observe that promised wages are almost always higher than paid wages and there appears to be relatively little relationship between wage promises and effort levels. On balance, the effect of firms being able to commit to real deferred incentives, i.e., the difference between FCT and the NCT control, is enormous. In other words, it is the comparison between FCT and NCT that really shows how successful Lazear's idea of deferment of wages is in practice.

B. Workers

Table 1 shows summary statistics of effort decisions following different types of wage offers. The table is separated into statistics for all rounds and for the last half of the each session, to capture the effects of subject learning in what is a fairly complex experiment. In this section, we compare the results from the FCT to those in the control treatments, the CFT and NCT. Discussion of the RT is reserved until the next section.

The first row shows the average worker earnings for the experiment. Earnings were very similar in the FCT and CFT, and much lower in the NCT. The next 6 rows of Table 1 show the distribution of effort levels in the different treatments, which uniquely maps into social surplus. As predicted by the model, the social surplus is higher in the FCT (averaging 132.9 per round) and CFT (137.9) than in the NCT (125.7). However, the gap is much less than predicted by the Lazear model (160 in the FCT versus 100 in the NCT). We return to this later.

The next two rows show the proportion of worker players playing the Lazear optimal, income-maximizing strategy in each stage. In each treatment a large majority of subjects play the optimal strategy. In the FCT, subjects play optimally low and optimally medium 77.8 and 61.1 percent of the time, respectively. In the CFT, these figures are 69.1 and 77.6 percent. In each treatment, a sizeable majority plays the optimal effort level at least 60 percent of the time (70 percent of subjects in the FCT and CFT and 73.3 percent in the NCT), and only a small minority of workers players systematically fail to optimize. This clearly demonstrates

that workers strongly respond to the monetary incentives provided. In other words, contract design has a strong influence on effort levels. This observation also supports the basic prediction of the Lazear model, namely that deferred compensation can be used to elicit increased effort.

The remaining rows examine deviations from the Lazear predictions in the FCT and CFT. There are more Fehr-Schmidt alpha rationalizable choices of low effort in the FCT than in the CFT, implying that social preferences provide an explanation for the greater prevalence of low effort in the FCT. There is also some evidence for reciprocity. Consider our earlier definition of reciprocity, where workers may wish to reward firms for wage offers beyond the alpha border (i.e., where $W_1 + W_2 + W_3 \geq 100$). While we find 18.7 percent supply high stage 1 effort in that region in the FCT, the corresponding number for the CFT is only 9.9 percent.

Finally, we split the region to the left of the Lazear line. Worker players were much more likely to choose medium effort in the region close to the line (where deferred pay is at least half of what is required to incentivize medium effort for a income-maximizing agent) than in the region close to the origin. This holds for both treatments and the difference between the two treatments is not significant. This suggests that for a fair proportion of the sub-optimally low responses, subjects understand that the amount of deferred payment is essential for triggering medium effort, but they appear to miscalculate the relationship between effort and payoffs.

To further understand workers' strategies, we consider a multinomial logit regression on their effort level in the first stage (0 = low, 1 = medium, 2 = high). In the first specification our independent variables of interest correspond to the axes in Figure 3, i.e. W_1 and (W_{12} + W_{13}).[16] In the second specification we add the Lazear line, its interactions with the pay variables, and a dummy for the Fehr-Schmidt triangle. We do not run the second specification for the NCT or the RT, as neither the Lazear line nor the Fehr-Schmidt region

[16] We use the notion W_{12} and W_{13} in the FCT and the CFT to denote the actual wage offer made in stage 1 for stages 2 and 3. In the NCT and RT it denotes the wage promise made in stage 1 for stages 2 and 3.

are properly defined in this treatment. As additional control variables we include ROUND, the number of times a subject has played the game, as a time trend which perhaps captures learning by doing; TEST SCORE, the score attained on a pre-experimental quiz, which perhaps captures understanding of the game or speed of calculation; and several personal controls obtained in a post-experiment questionnaire SEX, AGE, STUDY YEAR (undergraduate =1, 2, 3; masters = 4; PhD = 5, 6, 7), and ECONOMICS (1 if majoring in economics). We do not have expectations for the signs of the coefficients on the personal characteristic variables. To account for non-independence of observations, the regressions were clustered by subject and by session and were estimated using general linear latent and mixed models, GLLAMM (Sophia Rabe-Hesketh and Anders Skrondal, 2005).

The regression results for the wage variables are shown in Table 2.[17] We begin with some description of the main results from the FCT (columns 1 and 2). As predicted by the model, deferred pay has a large and statistically significant effect on effort levels. Stage 1 pay also has a statistically significant effect, though its magnitude is much smaller. The additional controls for the Lazear line and the alpha border are also strongly significant in the manner predicted by the theories. There is a discrete increase in the observed effort to the right of the Lazear line and considerably lower effort in the triangle where low effort is Fehr-Schmidt rationalizable. In addition, the interactions of the Lazear line and the pay variables are negative, suggesting the marginal effect of increasing pay on effort is lower in the region where medium effort is incentivized.

Next we compare the regression results to those in the CFT and the NCT. Worker players in the CFT responded to monetary incentives in a very similar manner as those in the FCT. The coefficients on deferred pay and the Lazear line are similar for the two treatments. However, as would be expected, the results on the variables capturing the role of social preferences are very different across the two treatments. As previously mentioned, W_1 is unambiguously a gift, as it is zero in the Lazear equilibrium. Unlike the case when workers play against a human firm player, increases in W_1 do not result in increased effort levels in the

[17] In the interest of saving space, we do not report results for the control variables. A full set of regression results is available from the authors on request.

CFT. In addition, in the CFT there is no significant difference between the frequency of low effort in the Fehr-Schmidt triangle and in the northeast region to the right of the Lazear line – again lending support to the role of social preferences when there are two real players.

Finally, the comparison between the FCT and the NCT suggests a strong role of credible incentives. In the NCT the coefficients on wage promises are close to zero and statistically insignificant, implying that, as predicted by the model, workers do not believe the promises to be credible and expect firms to renege. Thus it is impossible for firms to incentivize workers through deferred pay. However, the coefficients on the stage 1 offer are strongly significant and much larger than in the FCT. This is broadly consistent with the findings of Fehr, Kirchsteiger, and Riedl (1993), who in a one stage set-up find that subjects engage in non-enforceable reciprocal cooperation to bring about socially superior outcomes. However, the much smaller coefficients on the stage 1 offer in the FCT regression along with the importance of future wages, suggests that gift exchange plays a smaller role when it is possible to write legally enforcing contracts over multiple periods. In other words, formal and informal mechanisms that improve social efficiency appear to work as substitutes.

C. Firms

The behavior of firm players is more complex. Unlike worker players, whose decisions map directly into payoffs, firm players must anticipate the worker players' response to their wage offers. To the extent that worker players' actual strategies deviate from equilibrium and these deviations change over time, firm players must refine their beliefs about worker behavior and update their wage offers.

A second problem facing firm players is that the set up of the experimental game makes it difficult for them to come to a wage offer that incentivizes medium effort and is more profitable than the simple (0, 0, 0) offer. As pointed out already above, only wage profiles in the Fehr-Schmidt triangle can earn a firm more profit than (0, 0, 0). This implies two difficulties for firm players. First, the Fehr-Schmidt triangle is relatively small - it comprises only approximately 0.6 percent of the offer space. So there is essentially no chance

of finding it through luck. Firms need to reason about this. The second complication is that firms face a distribution of workers with different alphas. Some might have low or zero alphas and will exert medium effort for most wage profiles in the triangle; others will have higher alphas and may respond with low effort where another worker exerts medium effort. This makes learning about the optimal wage profile difficult. Moreover, notice that even (0, 0, 0) is sometimes met with medium effort due to errors by the worker players, which implies that the payoff gap between the optimal wage profile in the Fehr-Schmidt triangle and (0, 0, 0) is smaller than theory would suggest.

In order to better understand the firm's problem, we have calculated the firm player's expected payoff in the FCT under probabilistic response by the worker players using the regression results from the 1st column of Table 2 and a similar regression for stage 2.[18] In other words, we are trying to establish what would have been the profit-maximizing wage profile in the FCT treatment.

Under probabilistic response, the firm player's expected payoff is (we denote e = 0 if the round ends prior to a stage being reached):

$\sum_{e,t} \pi_{e,t} Z_{e,t} - W_t p_{e_{t-1,t-2}}$ where e = 0, L, M, H; t = 1, 2, 3

$p_{e_{t-1,t-2}}$ is the probability that the round has ended given effort in previous stages

$\pi_{e,t}$ is the probability that the worker player chooses effort e in stage t

We calculated the expected payoffs for all possible combinations of W_1, W_2, and W_3 in the FCT, and found that (0, 0, 0) gives an expected return of 107, whereas (0, 0, 60) gives 117. Given that (0, 0, 60) carries the downside risk associated with workers who have high alphas supplying low effort, it is perhaps not surprising that we observe so many worker players choosing (0, 0, 0) in the FCT.

[18] The regression for stage 2 differs slightly from that of stage 1 in that the independent variables for pay are (W_1 + W_2), past and present pay, and W_3, future pay. Regression results are available from the authors by request.

In order to further examine the firm players' choices, Table 3 shows some summary statistics of wage offers, which can be compared to the predictions of the model. As with the worker players, we have broken down the statistics into all 20 rounds and the second half of the experiment, to capture the effects of learning.

The first row shows firm profits. These were significantly higher in the NCT than in the FCT, implying substantial deviation from equilibrium behavior. The next two rows show the mean wage offers. Wage offers declined over the session in each treatment. Interestingly there is little difference between stage 1 offers in the FCT and NCT, with both averaging about 20 overall and 15 in the last 10 rounds. There is, however, a big difference in the actual offer ($W_2 + W_3$), with much higher wages in the FCT. When offers of (0, 0, 0) are excluded, deferred pay in the FCT averaged 65.4 for all 20 rounds and 58.2 for the last 10, which is very close to the Lazear equilibrium of 60. The next two rows show the shares of total pay in stage 1 and stage 3, excluding offers of (0, 0, 0), where shares are undefined. The model predicts that in the FCT the stage 1 share should go to zero and the stage 3 share should be at least two thirds. Neither of these predictions is fully supported by the data, stage 1 pay remains positive and stage 3 pay is well below two thirds. However, the data also show that firm players offer significantly more deferred pay in the FCT than in the NCT.

The last 6 rows provide summary statistics on the distribution of different types of wage offers. The equilibrium for both the FCT and NCT is $W_1 = 0$. In both treatments a sizable minority (which is increasing over the course of the experiment) follow this prediction, though the majority gave some sort of a gift in the first stage. In the FCT and NCT, respectively 30 percent and 33 percent of firm players make a positive first stage offer in all 20 rounds. The next row shows the percentage of subjects making a zero wage offer in the third stage. Consistent with the predictions of the Lazear model, positive wage third-stage offers are made in approximately 63 percent of observations in the FCT, but only 20 percent in the NCT. The next row shows the percentage offering (0, 0, 0), the equilibrium strategy for the NCT and a local maximum in the FCT. In both treatments, this strategy is employed by an increasing minority over the course of the experiment. However, in line with the predictions of the Lazear model, (0, 0, 0) offers are made significantly more frequently in the NCT than

the FCT. Finally, the final three rows show three measures of the extent to which the pay offers used deferred compensation to incentivize effort: the percentage of offers that incentivize medium effort in at least one stage, the percentage that incentivize medium effort in both stages, and the percentage that incentivize medium effort in both stages and provide the firm with higher expected payoff than (0, 0, 0). There are significant differences between the FCT and the NCT. In the FCT, a sizeable minority of firms pay deferred compensation; whereas in the NCT, where deferred compensation does not incentivize effort, almost none do so.

Overall the results imply that, despite the difficulty of the firm player's problem, the Lazear model has considerable predictive power for firm strategies. The comparison of offers across the three treatments is particularly illuminating, as firm players are able to incentivize workers with deferred compensation much more often in the FCT than in the NCT.

D. The Reputation Treatment

We kept our fourth treatment, the Reputation Treatment, RT, for separate discussion, as it differs from the other three in an important dimension. While we have clear-cut theoretical predictions for the main treatment FCT and its two controls, this is not the case for the RT.

The RT is like the NCT but with added information about the past. When workers are matched with a firm, they can access this firm's entire history. That is, they can see every past wage promise, the effort induced, and the actual wages paid. This allows firms to build up a reputation. In principle, such reputation building could be modelled formally (in the spirit of Kreps et. al., 1982) but not only would such computations be extremely cumbersome, they would also crucially depend on the assumptions made about behavioural types.

In fact, our motivation for including the RT is quite different from our motivation for the design of the FCT and its two controls. With the RT we are not so much interested in

testing a theory, rather we want to *explore* to which extent a more realistic setup – with reputation instead of commitment facilities – fares in incentivizing workers.

To get a first impression let us compare Figure 3D, for the RT, with Figure 3C, for the NCT. While there is little difference between the two on the x-axis where $W_1 = 0$, there is noticeably higher effort in the RT further to the north. It appears that wage promises that are backed up with some actual pay in the first period and, presumably, a good reputation for paying out workers in the past are far more successful than the mere (and often empty) promises of the NCT. In fact, towards the high northeast there are even several instances of high effort in the RT.

Comparing the RT to the FCT, we make two further interesting observations. Even with reputation building, promises are much higher than actual commitments and workers do not have full faith in these promises. Moreover, to the northeast of the alpha border there is much more low effort in the RT than in the FCT.

To further examine worker behaviour in the RT, we return to Tables 1 and 2. On balance, comparing reputation building with the extremes of full and no-commitment the figures suggest that the outcomes are somewhere in the middle. Table 1 shows that workers supply medium or high effort slightly more often in the RT than in the NCT. A comparison of the FCT and RT baseline regressions (columns 5 and 6) in Table 2 shows that workers increased their effort levels in response to pay promises in the RT, but not the NCT. When we add reputation variables in column 7, we find that workers supplied more effort when the firm had a history of making generous wage offers in the second stage. From the perspectives of Lazear's theory we believe this is positive news. As has been suggested in the literature, our data prove that reputation building can serve as a partial substitute for commitment.

Finally, we turn to the behaviour of the firm players. First we return to Table 3. Firm players offered slightly more deferred compensation in the RT than in the NCT, though still considerably less than in the FCT. Although Table 3 summarizes overall wage offers, it does not offer insights as to how the subjects came to those offers. The decision to pay deferred

wages in the NCT and RT comes only after the worker player has made at least one decision regarding effort. If firm players are reciprocal agents, it is likely that that the worker's effort level will have an effect on the firm's actually paid out wage in the later stages in both the NCT and RT. However, the two treatments differ in one crucial respect. In the RT, a firm player can build up a reputation for paying reciprocal wages, whereas in the NCT, reciprocal wages are a pure gift. A pattern of responding to increased effort with higher wages would be evident to future worker players, who might be more likely to respond to fair wage promises with higher effort. Because this effect is only present in the RT, it is likely that medium or high effort will trigger a greater wage response in this treatment than in the NCT. Table 4 shows regression estimates of the firm players' adaptive behavior in the NCT and RT. The regressions show that firm players increased their wage offers in response to past effort given by the workers, and that the extent of this increase is greater in the RT than in the NCT. This suggests that reciprocity works better if backed up by reputations as a mechanism to increase firms' wage offers and workers' effort levels.

As with the results for the worker players, the results for the firm players show that wage offers are higher and more likely to elicit effort in the RT than in the NCT, again suggesting that reputation building serves as a partial substitute for commitment.

IV. Conclusions

There exists a sizeable theoretical literature showing that deferred compensation can be used by firms to elicit higher effort from their employees. The effect of deferred compensation on worker effort has, however, thus far eluded direct empirical testing, as this sort of compensation structure is most likely to be used if it is impossible to measure all of the dimensions of short-term output and effort. Nevertheless, there has been indirect evidence for the relevance of deferred compensation. This paper supplements this indirect evidence by directly testing the relationship between the time structure of wages and worker effort in a laboratory experiment. It is the first experiment to investigate the role of the intertemporal allocation of wages for incentivizing workers.

We implement a simple three-stage game that retains the key characteristics of Lazear's model of deferred compensation. In our main experimental treatment we endow firms with full commitment possibilities such that the resulting subgame perfect equilibria exhibit the important characteristics of deferred compensation contracts. Crucially, pay is below productivity for young workers and above productivity for old workers.

At first sight, our data in the main treatment appear to be rather mixed. While there is evidence that deferred compensation works, there are also workers who are "Lazear incentivized" but shirk and there are many firms that do not make use of deferred wages. The picture becomes much clearer once we introduce two control treatments and examine the differences between the controls and the main treatment. The first control eliminates firms' commitment power such that it is impossible to offer credible deferred incentives. The result is an almost complete breakdown of worker-firm relations and a dramatic loss in efficiency compared to the main treatment. It is this comparison that really underlines the success of the Lazear model – the difference deferred compensation makes.

Our second control is designed to capture the effects of social preferences in our environment. By eliminating the firm subject and introducing computerized wage offers, social preferences become irrelevant in this setting. The contrast between this control and the main treatment highlights the importance of inequity aversion in our game. Worker subjects who in the Lazear equilibrium earn substantially less than firms have a tendency to reduce inequity by choosing low effort which harms the firm more than it harms them. There is, of course, substantial heterogeneity in workers' inequity aversion but we can identify a region of the wage profiles (the "Fehr-Schmidt triangle") where firms face the risk that Lazear optimal wages do not generate Lazear optimal effort. It is the characteristics of the Fehr-Schmidt triangle that make wage setting rather complicated for firms in our game which explains the wide variety of wage offers that we observe in all treatments.

There is also some limited evidence for reciprocity that transcends purely distributional concerns in our data. The most interesting finding on reciprocity is perhaps that reciprocity requires wages that are structured such that they cannot be understood as chosen

for pure incentivization purposes. Specifically, this means in our setup that firms can either choose a high first-stage wage which should be zero according to both the Lazear and Fehr-Schmidt models or need to choose wages that are exceed the wages required to incentivize even the most inequity averse worker.

Finally, we change tack and move away from theory testing to a rather more explorative and arguably more "realistic" treatment, the Reputation Treatment where firms have no commitment power but can build up a reputation for making credible promises for paying high wages when workers are older. We find that reputation building through actual payment of high wages to old workers is partially successful such that the Reputation Treatment achieves some of the efficiency gains that subjects reach in the ideal Full Commitment Treatment but not all. This indicates that deferred compensation can be expected to be effectively used whenever there are long-lived firms that can build a track record of paying rather than sacking older workers.

Figure 1
A Deferred Compensation Contract

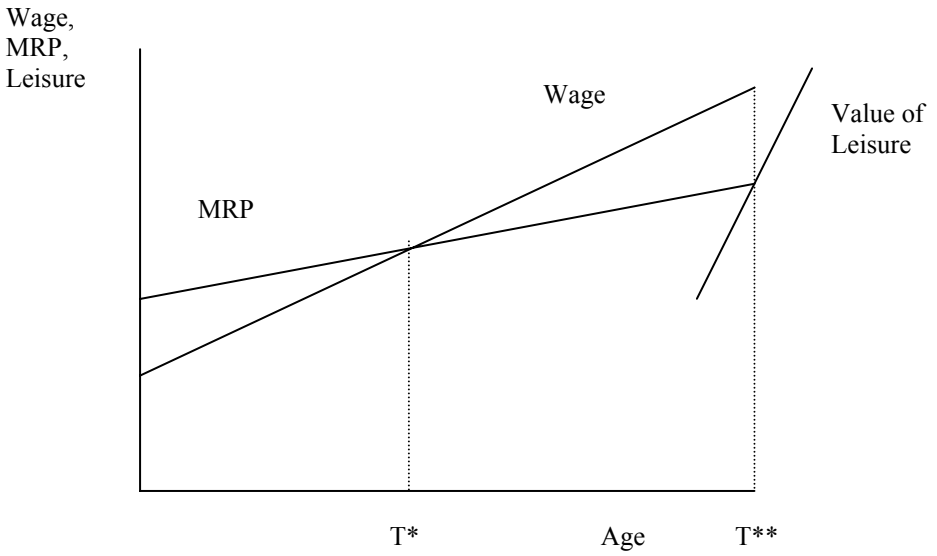


Figure 2
The Relationship Between Current Pay, Deferred Pay, and Effort

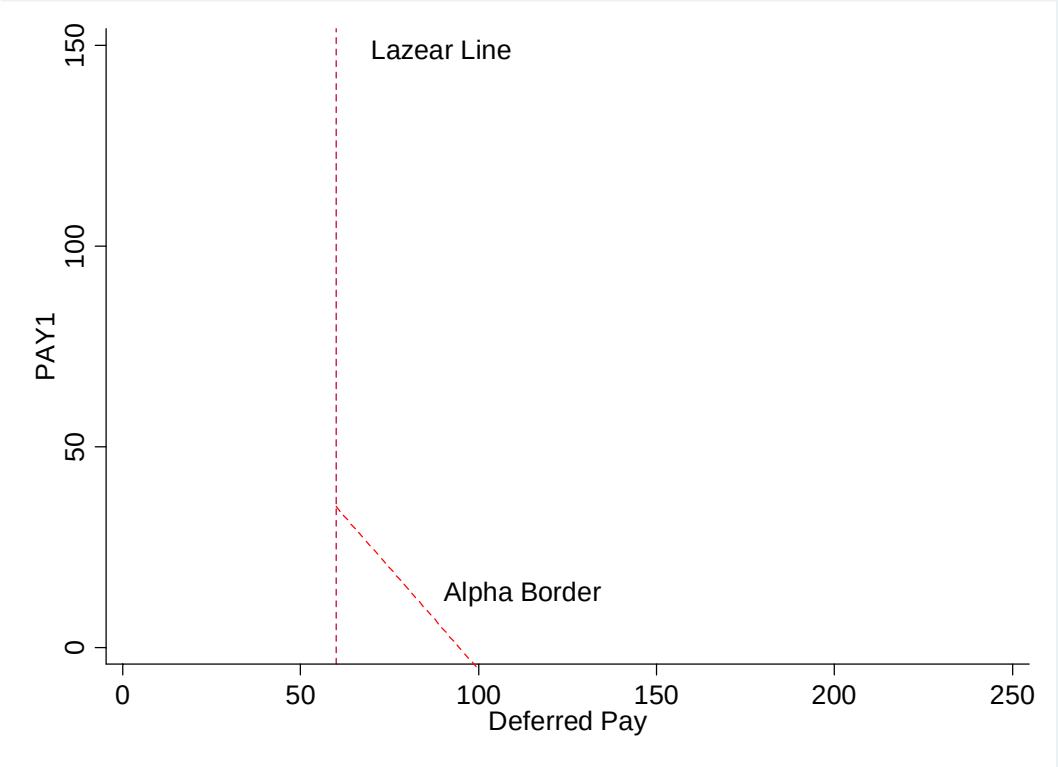
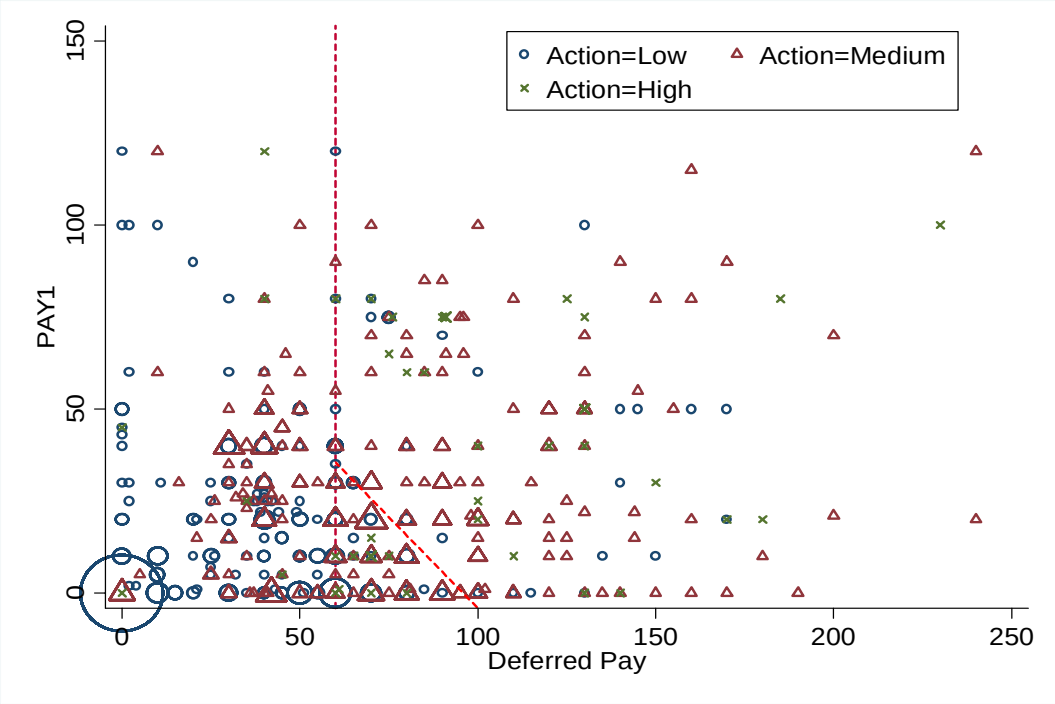
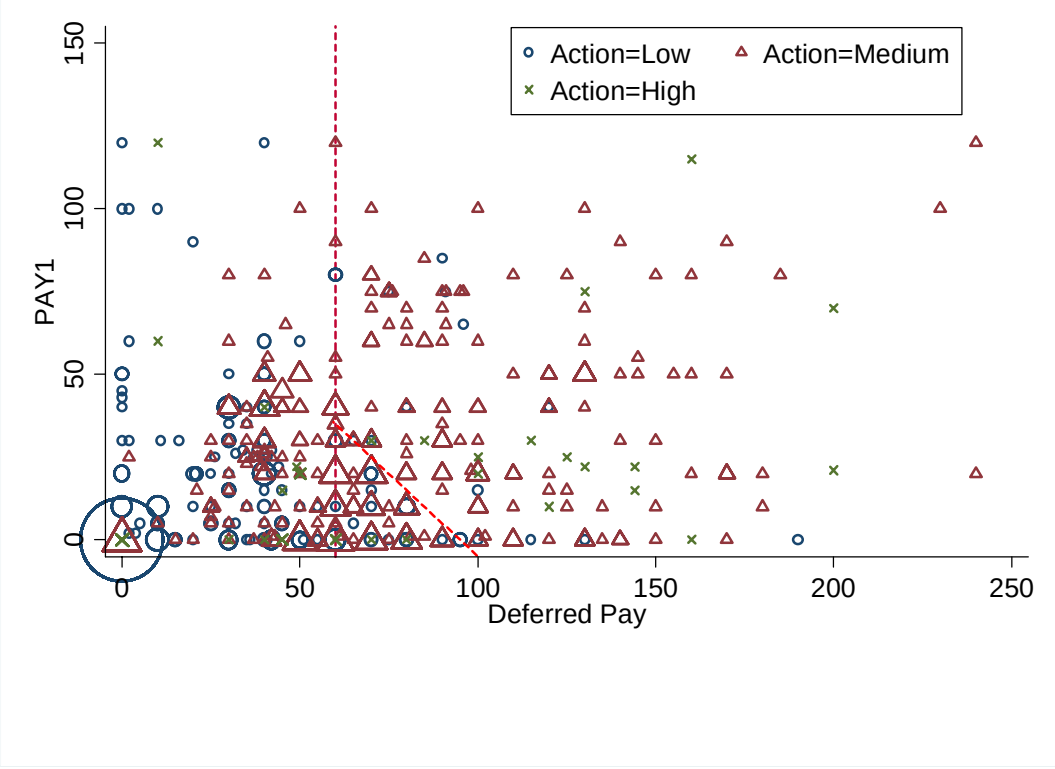


Figure 3
Current Pay, Deferred Pay, and Effort, Stage 1

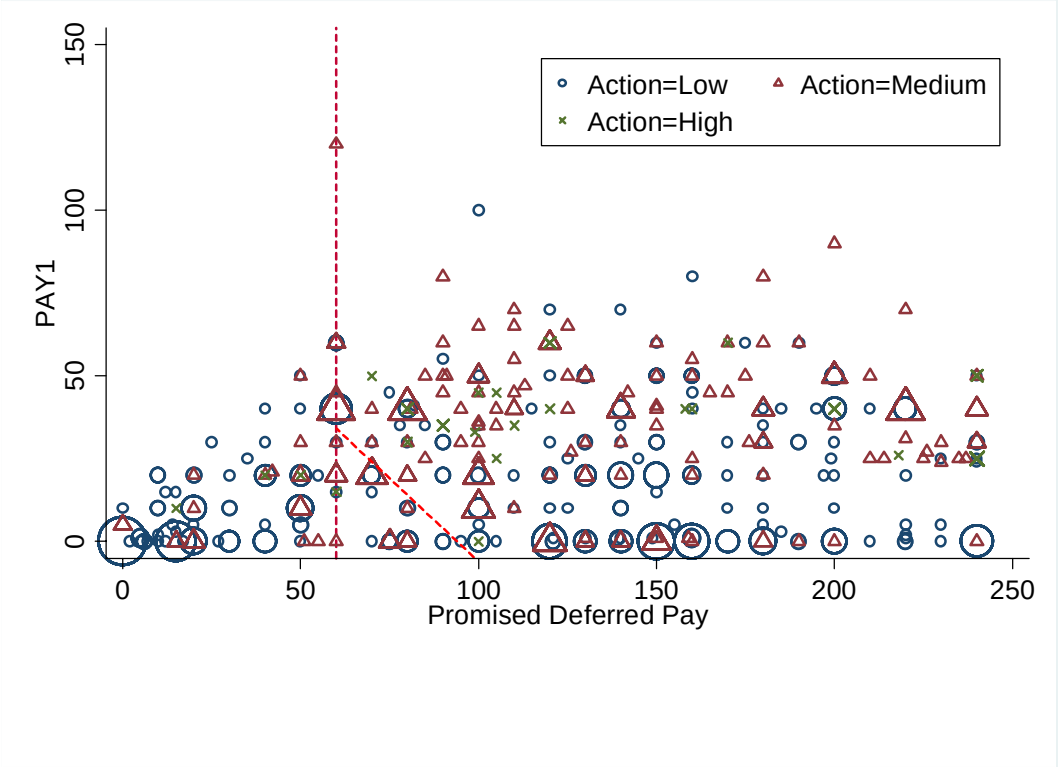


A. FCT

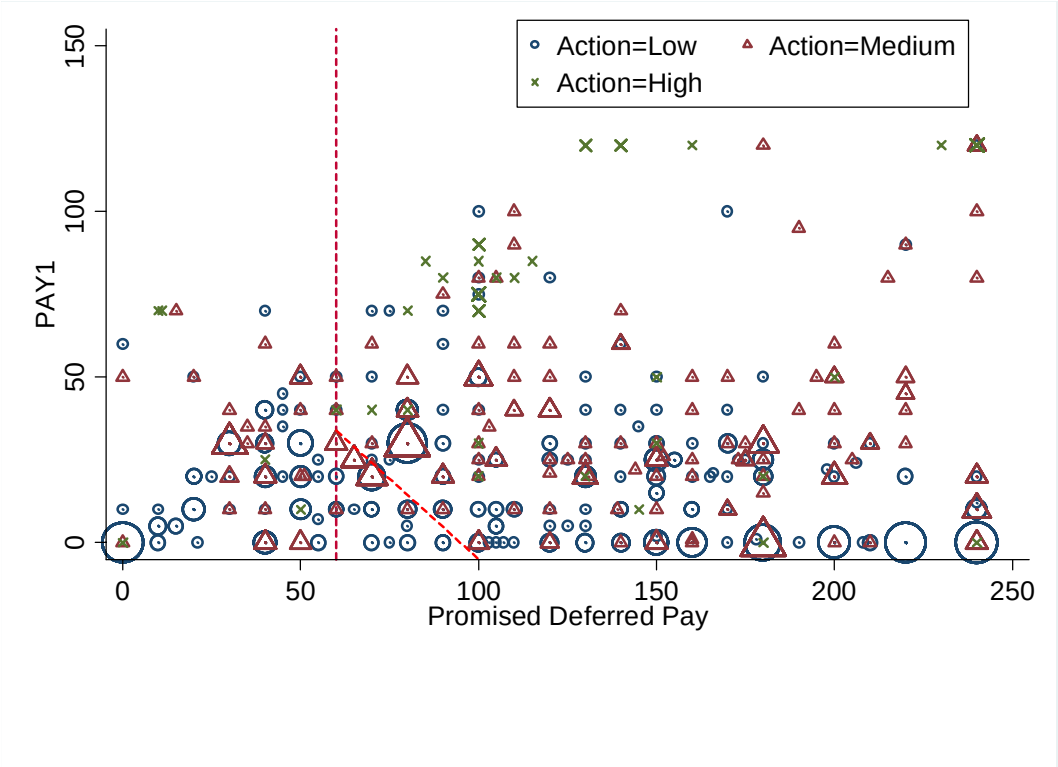


B. CFT

Figure 3, continued
Current Pay, Deferred Pay, and Effort, Stage 1



C. NCT



D. RT

Table 1
Summary Statistics on Worker Players' Effort Levels

		FCT	CFT	NCT	RT
Rounds 1-20	Mean Earnings (pounds)	17.32	17.29	12.93***	15.13
	E_1 = Low	49.8	40.5	60.5	56.7
	E_2 = Low	62.6	59.0	84.0***	76.6**
	E_1 = Medium	43.2	54.0	34.5	36.5
	E_2 = Medium	31.3	34.0	13.2***	19.7**
	E_1 = High	7.0	5.5*	5.0	6.8
	E_2 = High	6.1	6.9	2.8*	3.7
	Lazear optimal, stage 1	67.0	73.7	60.5	---
	Lazear optimal, stage 2	74.2	75.7	84.0**	---
	Fehr-Schmidt rationalizable, stage 1 (all obs)	7.5	4.7	---	---
	Fehr-Schmidt rationalizable, stage 1 (F-S region)	44.1	27.5	---	---
	E_1 = High if total offer >100	18.7	9.9*	---	---
	E_2 = High if total offer >100	13.2	10.7	---	---
	E_1 = Low if deferred pay < 30	89.6	82.2	---	---
	E_1 = Low if 30 $\leq$ deferred pay < 60	48.2	42.2	---	---
	E_2 = Low if deferred pay < 20	83.5	80.6	---	---
	E_2 = Low if 20 $\leq$ deferred pay < 40	55.4	55.6	---	---
Rounds 11-20	E_1 = Low	54.3	46.7*	68.0	67.3
	E_2 = Low	68.7	60.7	89.0**	86.0*
	E_1 = Medium	41.0	50.0*	27.3**	27.7**
	E_2 = Medium	28.7	34.5	8.4***	12.0***
	E_1 = High	4.7	3.3*	4.7	5.0
	E_2 = High	2.7	4.8**	2.6	2.0
	Lazear optimal, stage 1	71.7	75.3	68.0	---
	Lazear optimal, stage 2	79.3	75.9	89.0*	---
	Fehr-Schmidt rationalizable, stage 1 (all obs)	10.3	6.7*	---	---
	Fehr-Schmidt rationalizable, stage 1 (F-S region)	42.5	27.4	---	---
	E_1 = High if total offer >100	16.7	4.6	---	---
	E_2 = High if total offer >100	5.3	0.0	---	---
	E_1 = Low if deferred pay < 30	92.7	85.4	---	---
	E_1 = Low if 30 $\leq$ deferred pay < 60	49.4	50.6	---	---
	E_2 = Low if deferred pay < 20	88.2	79.6	---	---
	E_2 = Low if 20 $\leq$ deferred pay < 40	57.9	57.1	---	---

Notes: All figures except earnings are reported as percentages.
Stage 2 figures are only for observations where the coin toss after stage 1 was heads, and thus the round proceeded to stage 2 regardless of the worker's actions in stage 1.
*** = significantly different from the FCT at a 1% level
** = significantly different from the FCT at a 5% level
* = significantly different from the FCT at a 10% level
Significance calculated using a Mann-Whitney test. In order to ensure independence of observations, the session average is used as the unit of observation for the test (individual for CFT).

Table 2
GLLAMM Regressions on Worker Players' Actions in Stage 1

		FCT	FCT	CFT	CFT	NCT	RT	RT
E_1 = Medium	W_{11}	0.014**	0.017*	0.010	-0.001	0.046**	0.030**	0.027**
		(2.95)	(2.12)	(1.86)	(0.14)	(7.17)	(5.65)	(3.72)
	$W_{12} + W_{13}$	0.037**	0.053**	0.047**	0.062**	0.002	0.003*	0.002
		(9.69)	(4.83)	(10.14)	(5.48)	(0.95)	(1.97)	(1.28)
	Lazear line		2.87**		3.06**			
			(4.19)		(3.85)			
	Lazear line * W_{11}		-0.018		0.008			
			(1.67)		(0.64)			
	Lazear line *		-0.034**		-0.044**			
	($W_{12} + W_{13}$)		(2.83)		(3.34)			
	Fehr-Schmitt		-1.10**		-0.460			
	rationalizable		(2.69)		(0.99)			
	X's average W_{11}							-0.023**
								(2.64)
	X's average W_{22}							0.033**
								(3.43)
	X's average W_{33}							0.006
								(0.74)
	X's percent							0.014
	truthfulness							(1.89)
E_1 = High	W_{11}	0.029**	0.043**	0.009	0.023	0.059**	0.075**	0.078**
		(4.05)	(3.29)	(1.01)	(0.11)	(4.26)	(8.65)	(5.98)
	$W_{12} + W_{13}$	0.043**	0.044	0.048**	0.044*	0.004	-0.003	0.006
		(8.16)	(1.65)	(7.58)	(2.09)	(1.02)	(0.96)	(1.45)
	Lazear line		2.317		0.508			
			(1.75)		(0.40)			
	Lazear line * W_{11}		-0.026		-0.018			
			(1.59)		(0.96)			
	Lazear line *		-0.017		-0.010			
	($W_{12} + W_{13}$)		(0.60)		(0.43)			
	Fehr-Schmitt		0.084		0.219			
	rationalizable		(0.12)		(0.27)			
	X's average W_{11}							-0.043*
								(2.47)
	X's average W_{22}							0.058**
								(3.16)
	X's average W_{33}							-0.020
								(1.44)
	X's truth rate							0.023
								(1.95)
	Observations	600	600	600	600	600	600	557
	Log likelihood	-390.45	-379.18	-341.05	-324.68	-346.42	-426.34	-376.62

Notes: Absolute value of t-statistic in parentheses.
E_1 = low is the comparison group.
* and ** indicate significance at the 5 and 1 percent level, respectively.
The control variables are: SEX, AGE, YEAR, ECONOMICS, TEST SCORE, and ROUND.
The regressions include individual fixed effects and are clustered by session.

Table 3
Summary Statistics on Firm Players' Wage Offers

		FCT	NCT	RT
Rounds 1-20	Mean earnings (pounds)	16.23	18.81*	17.36
	$W_2 + W_3$ (pence)	54.74	21.31***	26.40***
	W_1 (pence)	21.73	19.89	20.76
	W_1/total offer	0.279	0.515***	0.473**
	W_3/total offer	0.316	0.080***	0.163**
	$W_1 = 0$ (percent)	35.8	31.9	33.3
	$W_3 = 0$ (percent)	36.8	80.0***	62.1**
	Wage offer = 0, 0, 0 (percent)	16.3	22.3*	20.5
	$W_3 \geq 40$ and/or $(W_2 + .5W_3) \geq 40$ (percent)	58.5	26.1**	24.2***
	$(W_2 + W_3) \geq 60$ and $W_3 \geq 40$ (percent)	21.8	3.0***	4.0**
	$(W_2 + W_3) \geq 60$, $W_3 \geq 40$ and $(W_1 + W_2 + W_3) \leq 100$ (percent)	12.8	0.6***	2.0**
Rounds 11-20	$W_2 + W_3$ (pence)	45.56	13.98**	21.71**
	W_1 (pence)	15.26	15.81	15.65
	W_1/total offer	0.239	0.545*	0.442**
	W_3/total offer	0.369	0.082**	0.192*
	$W_1 = 0$ (percent)	48.7	43.4	44.4
	$W_3 = 0$ (percent)	42.3	85.5***	68.1**
	Wage offer = 0, 0, 0 (percent)	21.7	32.5	30.6
	$W_3 \geq 40$ and/or $(W_2 + .5W_3) \geq 40$ (percent)	54.0	14.5***	18.1***
	$(W_2 + W_3) \geq 60$ and $W_3 \geq 40$ (percent)	19.3	2.4**	2.8**
	$(W_2 + W_3) \geq 60$, $W_3 \geq 40$ and $(W_1 + W_2 + W_3) \leq 100$ (percent)	18.0	1.2***	1.4***

Notes: * = different from the FCT mean at a 10% level
** = different from the FCT mean at a 5% level
*** = different from the FCT mean at a 1% level

Significance calculated using a Mann-Whitney test. In order to ensure independence of observations, the session average is used as the unit of observation for the test.

Figures for the NCT and RT are computed only for observations where the coin toss in both stages was heads, and thus the round proceeded to stage 3 regardless of offers or actions.

Table 4
Regression Results for Stage 2 and 3 Wage Offers

	NCT	NCT	RT	RT
Variable\Stage	2	3	2	3
W_{12}	0.050		0.134**	
	(1.33)		(2.57)	
W_{13}		0.038		0.015
		(1.63)		(0.65)
W_{23}		-0.054***		0.018
		(6.33)		(0.62)
E_1 = MEDIUM	11.66***	0.122	22.56***	4.93**
	(4.80)	(0.08)	(4.87)	(3.32)
E_2 = MEDIUM		7.18*		8.01**
		(2.21)		(3.04)
E_1 = HIGH	27.84**	-3.28	38.46***	0.946
	(4.03)	(1.17)	(9.00)	(0.32)
E_2 = HIGH		17.91		25.69*
		(1.92)		(2.44)
Constant	22.89***	3.43	13.14	7.39
	(5.20)	(0.70)	(2.20)	(1.29)
Observations	426	254	429	276
Adjusted R^2	0.265	0.071	0.346	0.209
F	10.00**	9.77**	448.4***	3.841*

Notes: The regressions include individual fixed effects and are clustered by session.
The reported R^2 is for the overall regression.
Absolute value of t-statistic in parentheses.
* indicates significance at a 10% level.
** indicates significance at a 5% level.
*** indicates significance at a 1% level.
The regressions control for ROUND.

References

Akerlof, George. "Labor Contracts as Partial Gift Exchange." Quarterly Journal of Economics, November 1982, 97 (4), pp. 543-569.

Brown, Martin; Falk, Armin; and Fehr, Ernst. "Relational Contracts and the Nature of Market Interactions." Econometrica, May 2004, 72 (3), pp. 747-780.

Brown, Martin; Falk, Armin; and Fehr, Ernst. "Competition and Relational Contracts: The Role of Unemployment as a Disciplinary Device." IZA Discussion Paper 3345, 2008.

Falk, Armin and Fischbacher, Urs. "A Theory of Reciprocity." Games and Economic Behavior, February 2006, 54(2), pp. 293-315.

Fehr, Ernst and Falk, Armin. "Wage Rigidity in a Competitive Incomplete Contract Market." Journal of Political Economy, February 1999, 107 (1), pp. 106-134.

Fehr, Ernst; Gachter, Simon; and Kirchsteiger, Georg. "Reciprocity as a Contract Enforcement Device: Experimental Evidence." Econometrica, July 1997, 65 (4), pp. 833-860.

Fehr, Ernst; Kirchsteiger, Georg; and Riedl, Arno. "Does Fairness Prevent Market Clearing?" Quarterly Journal of Economics, May 1993, 108 (2), pp. 437-460.

Fehr, Ernst; Kirchsteiger, Georg; and Riedl, Arno. "Involuntary Unemployment and Non-Compensating Wage Differentials in an Experimental Labour Market." Economic Journal, January 1996, 106 (434), pp. 106 – 121.

Fehr, Ernst; Schmidt, Klaus. "A Theory of Fairness, Competition, and Cooperation." The Quarterly Journal of Economics, August 1999, 114 (3), pp. 817 – 868.

Fischbacher, Urs. "2007 z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics, June 2007, 10(2), pp. 171-178.

Gokhale, Jagadeesh; Groshen, Erica; and Neumark, David. "Do Hostile Takeovers Reduce Extramarginal Salary Payments?" Review of Economics and Statistics, August 1995, 77 (3), pp. 470-85.

Güth, Werner; Schmittberger, Rolf; Schwarze, Bernd. "An Experimental Analysis of Ultimatum Bargaining." Journal of Economic Behavior and Organization, December 1982, 3 (4): 367—388.

Hutchens, Robert. "A Test of Lazear's Theory of Delayed Payment Contracts." Journal of Labor Economics, October 1987, 5 (4, part 2), pp. S153-S170.

Kreps, David; Milgrom, Paul; Roberts, John; and Wilson, Robert. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory, August 1982, 27 (2), pp. 245-52.

Lazear, Edward. "Why is there Mandatory Retirement?" Journal of Political Economy, December 1979, 87 (6), pp. 1261-64.

Lazear, Edward. "Performance Pay and Productivity." American Economic Review, December 2000, 90 (5), pp. 1346-1361.

Mauss, Marcel. The Gift: Forms and Functions of Exchange in Archaic Societies. Free Press: Glencoe, Il. 1954.

Prendergast, Canice. "The Provision of Incentives in Firms." Journal of Economic Literature, March 1999, 37 (1), pp. 7-63.

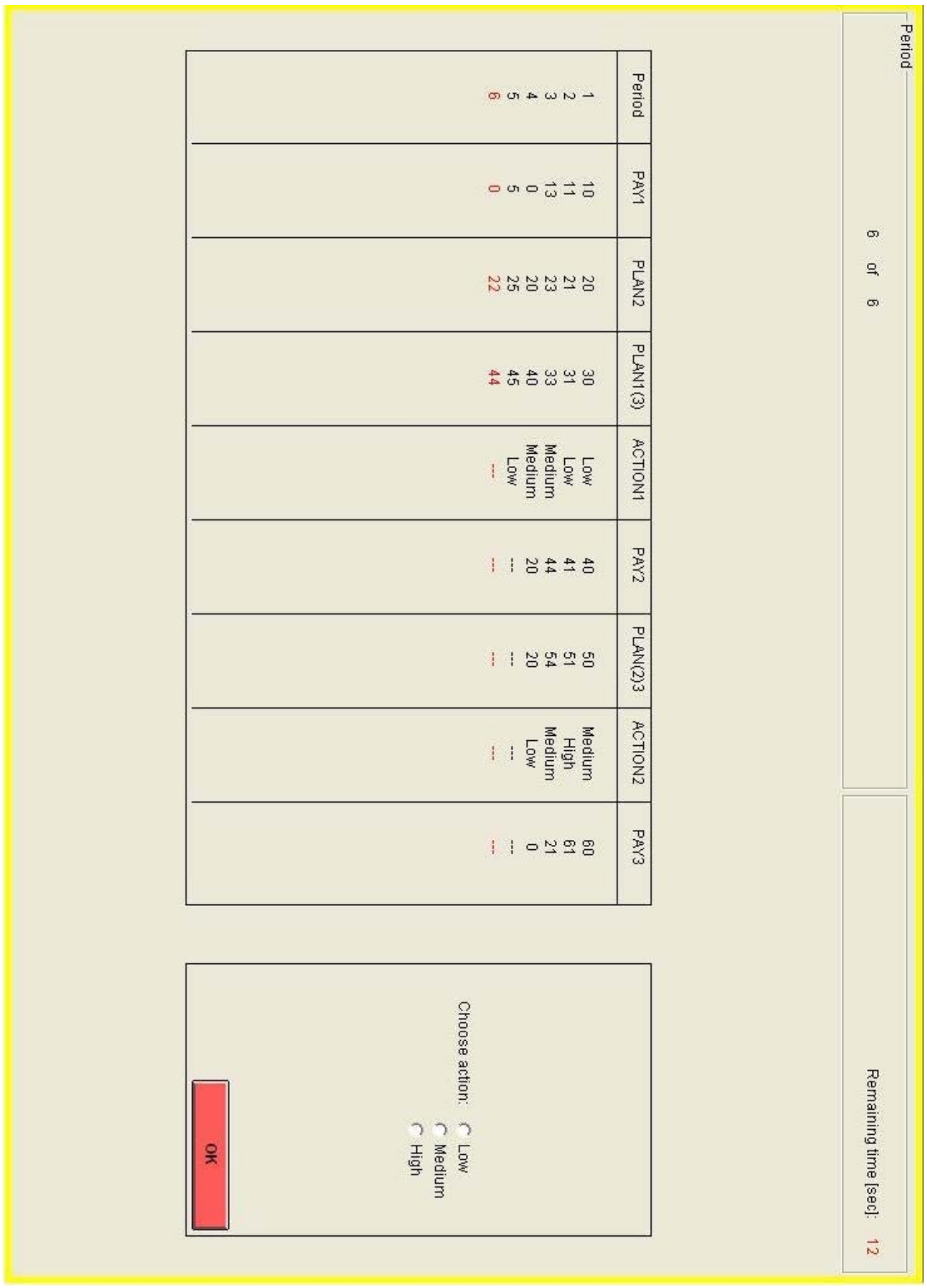
Rabe-Hesketh, Sophia and Skrondal, Anders. Multilevel and Longitudinal Modelling Using Stata. College Station, TX: Stata Press, 2005.

Salop, Steven and Salop, Joanne. "Self-Selection and Turnover in the Labor Market." Quarterly Journal of Economics, November 1976, 90 (4), pp. 619-27.

Seltzer, Andrew and Merrett, David. "Personnel Practices at the Union Bank of Australia: Panel Evidence from the 1887-1900 Entry Cohorts." Journal of Labor Economics, October 2000, 18 (4), pp. 573-613.

Not for Publication

Appendix 1: A Worker Player's Computer Screen from the RT



Not for Publication

Appendix 2: Instructions to Subjects

FCT Instructions

Welcome to our experiment!

Please be quiet during the entire experiment. Do not talk to your neighbours and do not try to look at their screens. Simply concentrate on what you have to do. If you have any questions, please raise your hand. We will come to you and answer it privately.

You are participating in an economics experiment in which you interact with other participants for twenty rounds. Depending on your choices, the other participants' choices and luck you can earn a considerable amount of money. The amount of money you will receive depends on how well you understand these instructions, so read carefully. You will receive the money immediately after the experiment.

In each round you will interact with one other participant who is chosen at random. Once a round is over, you will be matched with a new participant, again at random.

At the beginning of the experiment each participant is assigned one of two roles, either X or Y. Everybody keeps their role during the entire experiment.

Let us now describe what happens in each of the 20 rounds. At the beginning of the round X has to offer Y some money. More specifically, he has to decide about three different amounts that will be paid to Y at stages 1, 2, 3 of the round. Depending on the actions of Y, the round may end before stage 2 or 3, in which case the amount offered by X will NOT be paid for those stages. After X has made the offer of some money, Y has to take up to two decisions that determine the total earnings of both X and Y.

Sounds complicated? Don't worry. Here is a detailed description of the entire round.

At the start of each round X decides about three different amounts that he is willing to pay to Y. Let's call these amounts W1, W2, and W3. The value of each W1, W2, and W3 can be any amount between 0.00 and £1.20. This is the only decision that X will make during the round, but the cash payoff of that decision depends on the process described below, so everyone should read on carefully.

1st stage: Y is informed about W1, W2, and W3. Then he faces his own choice: He has to pick one of three different actions, which we call LOW, MEDIUM, and HIGH. Each of these three actions cause some costs for Y and some benefits for X. Let's call these amounts COST1 and BENEFIT1.

If Y chooses LOW, he faces COST1 of £0.00, while X earns BENEFIT1 of £0.50.
If Y chooses MEDIUM, he faces COST1 of £0.20, while X earns BENEFIT1 of £1.00.
If Y chooses HIGH, he faces COST1 of £0.40, while X earns BENEFIT1 of £1.40.

If Y chooses MEDIUM or HIGH he receives (£0.40 + W1 - COST1). The round then continues with stage 2. If Y has chosen LOW, chance decides whether the round continues or not. More specifically, the computer will flip a coin. If it lands on heads, the round continues to stage 2. If the coin lands on tails, the round is over. The final payoffs received by the two players for the round are as follows:

X receives (£1.00 - W1)
Y receives (£0.40 + W1 - COST1)

Both participants are informed about what happened so far and their earnings for the round.

2nd stage: Stage 2 is the same as stage 1., i.e., again Y decides whether to choose LOW, MEDIUM, or HIGH. This has the same payoff consequences as in stage 1 (i.e. COST2 and BENEFIT2 are the same as COST1 and BENEFIT1).

If Y chooses MEDIUM or HIGH he receives (W2 - COST2) in addition to his earnings from stage 1. The round then continues with stage 3. If Y has chosen LOW, the computer will flip a coin. If it lands on heads, the round continues to stage 3.

If the coin lands on tails, the round is over. The final payoffs received by the two players for the round are as follows:

X receives (BENEFIT1 + BENEFIT2 - W1 - W2)
Y receives (£0.40 + W1 + W2 - COST1 - COST2)

Both participants are informed about what happened so far and their earnings for the round.

3rd stage: There is no choice in stage 3 and what happens is very simple. X pays the amount W3 (which he has chosen in the very first stage) to Y. That is, if round 3 is reached, we will subtract W3 from X's total earnings in this round and will add W3 to Y's total earnings in this round.

The final payoffs received by the two players for the round are as follows:

X receives (BENEFIT1 + BENEFIT2 - W1 - W2 - W3)
Y receives (£0.40 + W1 + W2 + W3 - COST1 - COST2)

At the end of the 3rd stage both X and Y will be reminded of what happened in this round, i.e., they will see X's choice of W1, W2, and W3 as well as Y's choices between LOW, MEDIUM,

and HIGH. They will also see the final payoff for both X and Y in the round. Finally, you will see your total earnings so far, over all previous rounds.

At the end of the experiment you will be paid your total earnings over all 20 rounds in cash and be asked to sign a receipt.

This procedure may sound more complicated than it is. Take your time to read through the instructions again. And if you have any questions raise your hand. In a couple of minutes we will distribute an online questionnaire that we ask you to fill out. The purpose of this is simply to make sure that everybody fully understands the rules of the experiment before we actually start. You will need to answer each question correctly before beginning.

RT Instructions

Welcome to our experiment!

Please be quiet during the entire experiment. Do not talk to your neighbours and do not try to look at their screens. Simply concentrate on what you have to do. If you have any questions, please raise your hand. We will come to you and answer it privately.

You are participating in an economics experiment in which you interact with other participants for twenty rounds. Depending on your choices, the other participants' choices, and luck you can earn a considerable amount of money. The amount of money you will receive depends on how well you understand these instructions, so read carefully. You will receive the money immediately after the experiment.

In each round you will interact with one other participant who is chosen at random. Once a round is over, you will be matched with a new participant, again at random.

At the beginning of the experiment each participant is assigned one of two roles, either X or Y. Everybody keeps their role during the entire experiment.

Let us now describe what happens in each of the 20 rounds. At the beginning of the round X has to offer Y some money. More specifically, he has to decide about an amount that will be paid to Y at stage 1 of the round. He also has to inform Y about the amount that he plans to pay in stages 2 and 3. Depending on the actions of Y in stage 1, the round may end before stage 2 or 3, in which case the amount offered by X will NOT be paid for those stages. After X has made the offer of some money, Y has to take a decision that determines the total earnings of both X and Y. In the second stage X makes an offer of some money to Y and informs Y about

his plans for an amount to be paid in stage 3. These amounts need not necessarily be the same as X's initial planned amounts in stage 1. After X makes the offer, Y has to take a decision that determines the total earnings of both X and Y and whether the round will proceed to stage 3. In the third stage X makes an offer of some money to Y. This amount need not necessarily be the same as X's planned amounts in stages 1 and 2. Y has no further actions to take in this round.

Sounds complicated? Don't worry. Here is a detailed description of the entire round.

1st stage: At the start of this stage X decides about three different amounts that he is willing to pay to Y. Let's call these amounts PAY1, PLAN2, and PLAN3a. The value of each PAY1, PLAN2, and PLAN3a can be any amount between £0.00 and £1.20. X must pay Y PAY1, but he can later change his mind about PLAN2, and PLAN3a.

Y is informed about PAY1, PLAN2, and PLAN3a. These values will appear in red on the next screen. The screen will also show X's history throughout this session, that is all past values of PAY1, PLAN2, and PLAN3a offered by this particular X. Additionally, it will contain information about how the Y's responded to these offers (ACTION1 and ACTION2). This information will be in black. Y will not be able to determine the identity of the other Y players in previous rounds. The last page of this handout contains a picture of this screen.

Then Y faces his own choice: He has to pick one of three different actions, which we call LOW, MEDIUM, and HIGH. Each of these three actions cause some costs for Y and some benefits for X. Let's call these amounts COST1 and BENEFIT1.

If Y chooses LOW, he faces COST1 of £0.00, while X earns BENEFIT1 of £0.50.
If Y chooses MEDIUM, he faces COST1 of £0.20, while X earns BENEFIT1 of £1.00.
If Y chooses HIGH, he faces COST1 of £0.40, while X earns BENEFIT1 of £1.40.

If Y chooses MEDIUM or HIGH he receives (£0.40 + PAY1 - COST1). The round then continues with stage 2. If Y has chosen LOW, chance decides whether the round continues or not.

More specifically, the computer will flip a coin. If it lands on heads, the round continues to stage 2.

If the coin lands on tails, the round is over. The final payoffs received by the two participants for the round are as follows:

X receives (£1.00 - PAY1)
Y receives (£0.40 + PAY1 - COST1) where COST1 is zero because Y chose LOW at stage 1

Both participants are informed about what happened so far and their earnings for the round.

2nd stage: At the start of this stage X decides about two different amounts that he is willing to pay to Y. Let's call these amounts PAY2 and PLAN3b. The value of each PAY2 and PLAN3b can be any amount between £0.00 and £1.20. X must pay Y PAY2, but he can later change his mind about PLAN3b.

After X decides on PAY2 and PLAN3b Y is informed about these amounts on the screen described in stage 1 and shown on the back page of the instructions.

He then decides whether to choose LOW, MEDIUM, or HIGH. This has the same payoff consequences as in stage 1 (i.e. COST2 and BENEFIT2 depend on Y's choice in the same way as COST1 and BENEFIT1).

If Y chooses MEDIUM or HIGH he receives (PAY2 - COST2) in addition to his earnings from stage 1. The round then continues with stage 3. If Y has chosen LOW, the computer will flip a coin. If it lands on heads, the round continues to stage 3.

If the coin lands on tails, the round is over. The final payoffs received by the two participants for the round are as follows:

X receives (BENEFIT1 + BENEFIT2 - PAY1 - PAY2)
Y receives (£0.40 + PAY1 + PAY2 - COST1 - COST2) where COST2 is zero because Y chose LOW at stage 2

Both participants are informed about what happened so far and their earnings for the round.

3rd stage: X's actions in the third stage are similar to his actions in the second stage. At the beginning of the stage he must decide on the amount of money he is willing to pay to Y, PAY3. The value of PAY3 can be any amount between £0.00 and £1.20.

Y does not make any choices in stage 3 and what happens is very simple. X pays the amount PAY3 to Y. That is, if stage 3 is reached, we will subtract PAY3 from X's total earnings in this round and will add PAY3 to Y's total earnings in this round.

The final payoffs received by the two players for the round are as follows:

X receives (BENEFIT1 + BENEFIT2 - PAY1 - PAY2 - PAY3)
Y receives (£0.40 + PAY1 + PAY2 + PAY3 - COST1 - COST2)

At the end of the 3rd stage both X and Y will be reminded of what happened in this round, i.e., they will see X's choice of PAY1, PLAN2, PLAN3a, PAY2, PLAN3b, and PAY3 as well as Y's choices between LOW, MEDIUM, and HIGH. They will also see the final payoff for both X and Y in the round.

Finally, you will see your total earnings so far, over all previous rounds.

At the end of the experiment you will be paid your total earnings over all 20 rounds in cash and be asked to sign a receipt.

Below is the screen that the Y participant sees after X has made the offer of some money (i.e in stages 1 or 2). [The values chosen have been smudged out]

This procedure may sound more complicated than it is. Take your time to read through the instructions again. And if you have any questions raise your hand. In a couple of minutes we will distribute an online questionnaire that we ask you to fill out. The purpose of this is simply to make sure that everybody fully understands the rules of the experiment before we actually start. You will need to answer each question correctly before beginning. We will also distribute another questionnaire at the end of the experiment. This will ask some basic information about yourself and the approach you took during the experiment.

Below is the screen that the Y participant sees after X has made the offer of some money (i.e in stages 1 or 2). [The values chosen have been smudged out]