Econometrica, Vol. 75, No. 1 (January, 2007), 121–154

FAIRNESS AND CONTRACT DESIGN

BY ERNST FEHR, ALEXANDER KLEIN, AND KLAUS M. SCHMIDT[1]

We show experimentally that fairness concerns may have a decisive impact on the actual and optimal choice of contracts in a moral hazard context. Bonus contracts that offer a voluntary and unenforceable bonus for satisfactory performance provide powerful incentives and are superior to explicit incentive contracts when there are some fair-minded players, but trust contracts that pay a generous wage up front are less efficient than incentive contracts. The principals understand this and predominantly choose the bonus contracts. These results are consistent with recently developed theories of fairness, which offer important new insights into the interaction of contract choices, fairness, and incentives.

KEYWORDS: Moral hazard, incentives, bonus contract, trust contract, fairness, inequity aversion.

1. INTRODUCTION

THIS PAPER ADDRESSES the question of how concerns for fairness affect the actual and the optimal choice of contracts. We conducted a series of experiments where principals could choose which type of contract to offer to the agents. The optimal type of contract according to standard contract theory proves to be far less efficient than this theory predicts, while contracts predicted to be very inefficient if all agents are purely self-interested turn out to be superior. The experimental results suggest that this reversal in contract efficiency is due to the existence of fair subjects, because they exert a decisive impact on the incentive properties of different types of contracts. The principals in our experiments seem to understand this quite well, and a large majority indeed choose a contract that relies on fairness as an enforcement device. Those who choose the contract predicted by standard contract theory do very poorly. In the fi-

[1]A previous version of this paper was circulated under the title "Contracts, Fairness, and Incentives." We would like to thank three anonymous referees and a co-editor for very useful comments and suggestions. In addition, Sam Bowles, Colin Camerer, Mathias Erlei, Dan Friedman, Gilat Levy, Paul Milgrom, Massimo Motta, Larry Samuelson, Christina Strassmair, and Jean Tirole and seminar participants at the California Institute of Technology, the University of California at Berkeley and at Santa Cruz, the Harvard–MIT organizations workshop, Bocconi, Princeton, Stanford, Yale, the Universities of Copenhagen, Maryland, Pennsylvania, Toulouse, Washington, and Wisconsin (Madison), the University College London, and the European University Institute made helpful comments and suggestions. Susanne Kremhelmer and Hannah Hörisch provided excellent research assistance for the experiments discussed in this paper. Part of this research was conducted while the third author visited Stanford University and he would like to thank the Economics Department for its great hospitality. Financial support by Deutsche Forschungsgemeinschaft through SFB-TR 15 and Grant SCHM-1196/4-1, and the RTN network ENABLE is gratefully acknowledged. Ernst Fehr also gratefully acknowledges support from the Swiss National Science Foundation (project number 101312-103898/1) and the Research Priority Program on the "Foundations of Human Social Behavior" at the University of Zurich.

nal part of the paper, we show that these results are broadly consistent with a simple model of fairness.

To better understand the nature of our results, consider one of our experiments in more detail. Suppose that the principal wants to induce the agent to expend effort that is personally costly to her. Both parties can observe effort, but the courts can only verify it if the principal invests in a verification technology. If the principal makes this investment, she can offer an "incentive contract" to the agent, which fines the agent for unsatisfactory performance. The problem with the explicit incentive contract is that the verification technology is imperfect and the fine that can be imposed on the agent is limited, meaning that the highest effort level that can be implemented is positive but falls short of the efficient level of effort. Alternatively, the principal can offer a "bonus contract" that does not rely on effort verification and enforcement by third parties. Instead, the principal announces a nonbinding, voluntary bonus payment if the agent's effort is satisfactory. This bonus contract is an implicit contract because third parties do not enforce the principal's bonus payment.

Given that each principal interacts with each agent only once in the experiment, a selfish principal would never pay the bonus. If it were common knowledge that all principals are selfish, rational agents would choose the minimum effort level. Thus, standard contract theory forecasts that the bonus contract is doomed to fail, while predicting that the incentive contract will do much better. Yet, the overwhelming majority of principals offered a bonus contract in our experiments. Even though many principals did not pay the bonus, a substantial number of them made quite generous bonus payments, inducing the agents to spend much more effort than under an incentive contract. Thus, the bonus contract induces more efficient effort choices and, therefore, the principals predominantly prefer the bonus contract relative to the incentive contract. Our experiments indicate that the principals' contract choices differ from those predicted by the self-interest model because concerns for fairness strongly affect the incentive properties of the contracts.

However, we also conducted a second experiment in which the self-interest model did quite well. In this experiment the principal was restricted to the choice between an incentive contract and a "trust contract." A trust contract offers a (generous) fixed wage to the agent and asks him to return this favor by choosing a high effort level. Standard contract theory again predicts that principals will choose the incentive contract. While the bonus contract appeals to the fairness of the principal to reward high effort, the trust contract appeals to the fairness of the agent to reciprocate a generous fixed wage. The trust contract did rather poorly, however, in contrast to the bonus contract. Many principals experimented with the trust contract, but they incurred losses on average and eventually most of them shifted to the superior incentive contract.

Thus, our experimental evidence not only contrasts to the viewpoint of standard contract theory, but also constitutes a challenge for theories of fairness. Why does the incentive contract outperform the trust contract—as standard contract theory predicts—while the bonus contract surpasses the incentive contract, contradicting standard theory? How can the remarkable performance difference between the trust and the bonus contract be explained when, after all, both contracts rely on fairness as an enforcement device? How can the poor performance of trust contracts be reconciled with the experimental results of Fehr, Kirchsteiger, and Riedl (1993) that showed that the trust contract would do fairly well?

To address these questions, we provide a unified interpretation of our results in terms of a simple model of fairness—the Fehr–Schmidt (1999) model of inequity aversion—in the final part of the paper. We primarily chose this model because of its tractability. Inequity aversion is a simple extension of the standard self-interest model that takes into account the fact that some people are not only interested in their own material payoff, but also dislike inequity. The model implies that the incentive contract, which is optimal when all actors are purely self-interested, is much less efficient when a share of people cares about fairness. Furthermore, bonus contracts that would be very inefficient if all actors were selfish are predicted to achieve much higher levels of efficiency when there are some fair-minded people. The model also shows, however, that the presence of some fair-minded people alone does not suffice to implement efficient behavior. The incentive structure of the "gift exchange" is also very important. The trust contract makes trusting behavior more risky than does the bonus contract. The model shows that the superiority of the bonus over the trust contract holds for general valuation and cost functions, and is thus not restricted to our experimental parameters. As a general principle, the person who has less at stake from trusting another should be the first to trust. We also show how the performance of the trust contract depends on the players' payoff functions. The model predicts that the trust contract fails for the payoff functions used in our experiments, but that it is profitable to offer trust contracts in the case of the payoff functions implemented by Fehr, Kirchsteiger, and Riedl (FKR) (1993). Thus, the model's major predictions are consistent with the observed qualitative pattern of contract choices in our and in previous experiments. In addition, the model makes some surprisingly accurate quantitative predictions of average behavior.

A large number of empirical papers have examined the effectiveness of different incentives schemes over the last 10–15 years. This literature was surveyed in Prendergast (1999) and in Chiappori and Salanié (2003). Both survey papers conclude that incentives matter, i.e., agents often seem to respond to changes in incentives in ways that are consistent with the predictions of prevailing principal–agent models. However, both papers also report that the evidence for the predictions of contract choices is much weaker. This is a main reason why we focused on the principals' contract choices between different types of contracts in our experiments. In this way, our experiments may contribute to a better understanding of the forces that determine which contracts prevail.

Chiappori and Salanié (2003) emphasized that problems of unobserved heterogeneity and endogenous selection often complicate clean inferences about the incentive effects of contracts. In fact, these problems result in an ambiguous interpretation of correlations between different contracts and different behaviors. Do the contracts induce the corresponding behaviors or are the behavioral differences across contracts the result of self-selection of heterogeneous individuals to different contracts? This problem is, in our view, particularly severe in the context of fairness preferences because there is little hope that nonexperimental field data allow the control of self-selection according to such preferences. Therefore, an experimental approach to these questions can offer additional insights. For example, we can control self-selection of agents to contracts in our experiments because principals and agents were randomly matched. Therefore, we observe the behavior of the same agents under different contracts. In addition, we can also control the exact effort cost function of the agent and the confounding effects of repeated interactions. All of this would be much more difficult, if not impossible, in the field.

Previous work by Camerer and Weigelt (1988), Fehr, Kirchsteiger, and Riedl (1993), Berg, Dickhaut, and McCabe (1995), Fehr, Gächter, and Kirchsteiger (1997), Hannan, Kagel, and Moser (2002), and Charness (2004) indicates that fairness concerns may play an important role in moral hazard contexts. However, these papers did not study how the principals choose between explicit and implicit incentives. Several experimental studies in the past few years have examined how the provision of explicit incentives affects the agents' behavior in a moral hazard context.[2] Many of these studies find indications that concerns for fairness and reciprocity affect the acceptance of explicit incentive contracts. However, the principals could not choose between *different types of contracts* in these papers—in particular, between explicit incentive contracts and implicit bonus contracts. Precisely this setting enables us to identify the strengths and the limits of the standard approach in contract theory by isolating conditions under which the model's contract choice predictions are met and conditions under which these predictions fail.

The rest of this paper is organized as follows. Section 2 discusses the principal–agent problem and the contractual possibilities. Section 3 describes

[2]DeJong, Forsythe, Lundholm, and Uecker (1985) showed how different institutional remedies, such as liability rules, mitigate the moral hazard problem. Schotter, Bull, and Weigelt (1987) studied the effects of piece rates and tournament incentives, and Schotter and Nalbantian (1997) examined the performance of various group incentive schemes. Chaudhuri (1998) investigated the ratchet effect in a dynamic principal–agent experiment in which the principals chose output contingent wages. Likewise, Güth, Klose, Königstein, and Schwalbach (1998) examined a multiperiod principal–agent game in which the principals could offer linear profit-sharing contracts. Cooper, Kagel, Lo, and Gu (1999) studied how Chinese students and managers respond to the incentives underlying the ratchet effect. Keser and Willinger (2000), Güth, Königstein, Kovacs, and Zala-Mezo (2001), and Anderhub, Gächter, and Königstein (2002) also studied the performance of output-contingent wages in a moral hazard context.

the experimental design and procedures. Section 4 reports the results of our main experiments and of a control experiment, showing that our results are not affected by framing effects. We offer a theoretical interpretation of the experimental results in Section 5. Section 6 summarizes our main results and concludes. An appendix that contains a theoretical analysis of Section 5 and some additional experimental data is provided on the supplementary material web site (Fehr, Klein, and Schmidt (2007a, 2007b)).

2. A SIMPLE PRINCIPAL–AGENT PROBLEM

Consider a principal who hires an agent to carry out production. If the agent expends effort $e \geq \underline{e}$, he generates a gross profit $v(e)$ for the principal that is strictly increasing and concave in e, but he also incurs a private cost $c(e)$ (measured in monetary terms with $c(\underline{e}) = 0$, $c'(e) > 0$, and $c''(e) > 0$). Let $e^{\text{FB}} > \underline{e}$ denote the unique first best efficient effort level that maximizes $v(e) - c(e)$.

Gross profits and effort costs cannot be contracted upon. Both parties observe the agent's effort level, but to contract on effort, it has to be verified by the courts. At date 0, before the agent chooses e, the principal can invest in a verification technology at a fixed cost k that permits partial verification of effort. To fix ideas, we assume that if the principal invested k and required the agent to work at least e^*, then with probability p, $0 < p < 1$, the courts observe whether $e \geq e^*$ or $e < e^*$. The principal can impose a fine f on the agent if shirking ($e < e^*$) had been verified. However, the agent cannot be punished arbitrarily harshly, i.e., the fine f is bounded above by $\bar{f}$. Let $\hat{e}$ denote the highest effort level such that $p\bar{f} \geq c(e)$, i.e., $\hat{e}$ is the highest effort level such that it is more profitable for a risk neutral agent to choose this effort level than to shirk (choose $e = \underline{e}$) and to incur the expected punishment $p \cdot \bar{f}$. We will call $\hat{e}$ the *highest incentive-compatible effort level*. To make the problem interesting, we assume that $v(\hat{e}) - c(\hat{e}) - k > v(\underline{e}) - c(\underline{e})$ but that $\hat{e} < e^{\text{FB}}$.

The timing of events is as follows. At date 0, the principal decides whether to incur the verification cost and offers a take-it-or-leave-it contract to the agent. If the agent rejects the offer, both parties get their reservation utilities that are normalized to 0. If the agent accepts, he has to choose e at date 1. At date 2, a random draw determines whether the agent's effort is verifiable (in case k has been invested). Then payoffs are realized and payments are made.

In this contractual environment we focus on three different types of incentive mechanisms that the principal may use:

1. *Incentive Contract* (IC): The principal invests in the verification technology and offers a contract (w, e^*, f) that stipulates a wage w, a demanded effort level e^*, and a fine f, to be paid in case shirking ($e < e^*$) is verified. The agent's (expected) monetary payoff with an IC is given by $M^A = w - c(e)$ if $e \geq e^*$ and by $M^A = w - c(e) - pf$ if the agent shirked ($e < e^*$). The principal's expected monetary payoff is defined by

$M^P = v(e) - w - k$ in case of $e \geq e^*$ and by $M^P = v(e) - w + pf - k$ if the agent shirked.

2. *Trust Contract* (TC): The principal does not invest in the verification technology and offers the agent an *unconditional* payment w (that may be generous). In return, she asks the agent to put in effort $e^* > \underline{e}$. However, if the agent accepts a trust contract, he cannot be forced to choose $e = e^*$. The monetary payoff from a trust contract (w, e^*) is given by $M^A = w - c(e)$ for the agent and $M^P = v(e) - w$ for the principal, where e is the agent's actual effort level.
3. *Bonus Contract* (BC): The principal offers an unconditional base wage w and asks the agent to expend effort $e^* > \underline{e}$. Furthermore, the principal announces her intention to pay a bonus b^* if the agent chooses $e \geq e^*$. However, neither the agent's effort nor the principal's bonus payment is enforceable. If the agent accepts a bonus contract, he chooses his actual effort level e at date 1. Then, at date 2, the principal is informed about e and chooses the actual bonus b. The principal is not obliged to pay $b = b^*$ but can choose any $b \geq 0$. A bonus contract implies monetary payoffs $M^A = w - c(e) + b$ for the agent and $M^P = v(e) - w - b$ for the principal.

From the point of view of standard contract theory, the analysis of the (second best) optimal contract is straightforward if we make the traditional assumption that the principal and the agent both want to maximize their material payoffs. If the principal chooses an incentive contract, he should use the maximum fine ($f = \bar{f}$), require the agent to choose the highest incentive-compatible effort level ($e^* = \hat{e}$), and offer a wage that holds the agent down to his reservation utility of 0 ($w = c(\hat{e})$). This contract induces the agent to choose $e = \hat{e}$ and yields monetary payoff $M^P = v(\hat{e}) - c(\hat{e}) - k$ for the principal. On the other hand, both the trust contract and the bonus contract are obviously doomed to fail, because e^* and b^* are "cheap talk" and cannot be enforced: With a trust contract the agent knows that his wage is fixed independently of his effort level, so he will always choose $e = \underline{e}$. With a bonus contract a self-interested principal will never pay a bonus, independent of the agent's effort choice. Anticipating this, the agent will again choose $e = \underline{e}$. Thus, the self-interest model predicts that the IC will dominate both TC and BC.

However, this analysis rests on the important assumption that both players are only interested in their own material payoffs. If principals and agents are also motivated by concerns for fairness and reciprocity, the outcome is less clear. By offering a generous trust contract, the principal can appeal to the fairness of the agent, and the agent may indeed reciprocate by providing $e > \underline{e}$. If the agent is offered a bonus contract, he may choose a high effort level so as to appeal to the fairness of the principal, and the principal may indeed reciprocate by paying a bonus voluntarily. Thus, both TC and BC may be more efficient than the self-interest model predicts. Such a change in the relative efficiency of the different contracts may then induce the principals to prefer a

TC or a BC over an IC. The question of whether TC and BC are more efficient than IC remains open, however, and cannot be answered on the basis of general, qualitative notions of fairness or in the absence of empirical evidence. Therefore, we implemented a series of experiments in which the principals could choose whether to offer an incentive, a trust, or a bonus contract.

3. EXPERIMENT DESIGN AND PROCEDURES

Tables I and II summarize important experimental details across the different contract types. The agents could choose effort $e \in \{1, \dots, 10\}$ with effort costs given by Table I. An effort of e yields a gross profit $v(e) = 10 \cdot e$ to the principal. If the principal invests in the verification technology at cost $k = 10$, she can verify the agent's effort with probability $p = 1/3$. The maximum fine the agent can be charged is bounded above by $\bar{f} = 13$. Note that in a first best world, the total surplus would be maximized if the principal did not invest in verification and the agent chose $e = 10$, which would yield a total surplus of $v(e) - c(e) = 80$. The principal is constrained to choose $w \geq c(e^*)$ in all types of contracts, which rules out losses for the agent if he meets the principal's effort demand. We imposed this constraint to ensure that loss aversion does not affect the agents' behavior.

Given these parameters, a self-interested agent who maximizes his expected payoff can be induced to choose an effort level of at most 4 by imposing the maximum fine of 13. Thus, if both parties are self-interested, the optimal incentive contract stipulates $f = 13$, $e^* = 4$, and $w = 4$, which limits the agent to his reservation utility. In equilibrium, the monetary payoffs are $M^A = 0$ and $M^P = 26$. The self-interest model also predicts that the agent always chooses $e = 1$ under a trust or a bonus contract. Therefore, the principal would offer $w = 0$ and monetary payoffs are $M^A = 0$ and $M^P = 10$ for these contracts.[3]

The experimental subjects were students of the University of Munich and the Technical University of Munich. We had 20–24 subjects in each session, half of them randomly assigned to the role of the principal and half to that of the agent. The two groups were located in separate rooms. All subjects had to read detailed instructions and to solve several exercises before the experiment started to ensure that all of them understood the rules of the experiment.

TABLE I

THE AGENT'S EFFORT COST FUNCTION

e	1	2	3	4	5	6	7	8	9	10
$c(e)$	0	1	2	4	6	8	10	13	16	20

[3]Note that the agent is indifferent whether to accept or to reject this contract. Because wages are discrete, a second equilibrium exists in which the principal offers a wage that is one token higher. This increases the agent's payoff while decreasing the principal's payoff by 1.

TABLE II

TIME STRUCTURE, FEASIBLE CONTRACTS, AND PAYOFFS

	Incentive Contract	Trust Contract	Bonus Contract
Date 0	P offers (w, e^*, f)	P offers (w, e^*)	P offers (w, e^*, b^*)
Date 1	A chooses e	A chooses e	A chooses e
Date 2	If $e < e^*$ nature determines whether fine has to be paid	—	P chooses b
Expected monetary payoffs	$M^P = 10e - w - k + \begin{cases} 1/3 \cdot f, & \text{if } e < e^* \\ 0, & \text{if } e \geq e^* \end{cases}$	$M^P = 10e - w$	$M^P = 10e - w - b$
	$M^A = w - c(e) - \begin{cases} 1/3 \cdot f, & \text{if } e < e^* \\ 0, & \text{if } e \geq e^* \end{cases}$	$M^A = w - c(e)$	$M^A = w - c(e) + b$

We had 10 periods in each session. The agents were randomly matched with a different principal in each period. The randomization procedure ensured that no agent interacted more than once with the same principal. Thus, we had 10 contracts with 10 different contracting partners for each subject in each experimental session.

After each period, the subjects had to compute their own payoff and that of their partner. The outcome of each period remained strictly confidential to rule out the possibility of reputation building, that is, each principal–agent pair observed only what happened in their own relationship. They did not observe the contracts offered by the other subjects in the room; neither did they observe their current partner's past behavior. Furthermore, the matching was random and anonymous, i.e., the subjects' identity was never revealed to the other players. Finally, the subjects collected their total monetary payoffs privately and anonymously at the end of the session.[4] Each subject received an initial endowment of €10 ($\approx$ U.S. \$12.5 at the time of the experiment). The experimental (token) payoffs were converted into real money at the rate of 1 token = €0.1. The average payoff was about €22.5 (U.S. \$28).[5] Each session lasted between 2 and $2\frac{1}{2}$ hours. A complete set of the instructions for all our experiments can be found on our web page.[6]

[4]The students were told at the beginning of the experiment that they will leave the laboratory after the experiment individually one after the other and that they will receive their payment in private. This is exactly what happened.

[5]The highest total income was €57 (U.S. \$72), corresponding to an hourly wage of roughly €25 (U.S. \$31). However, the subjects could also incur substantial losses. To avoid the possibility of having a subject finish with negative earnings, he was expelled from the experiment if his accumulated earnings fell below €2.50 (U.S. \$3.12); this occurred only three times.

[6]The experimental instructions (translated into English) and the data of our experiments are available on the *Econometrica Supplementary Materials* web site (Fehr, Klein, and Schmidt (2007b, 2007c)).

We conducted six experimental sessions, each with different subjects. In sessions S1 and S2 we studied how the existence of fairness concerns affects the principals' choice between TC and IC by implementing the trust-incentive (TI) treatment, in which only a TC or an IC could be offered to the agents. In a second step (sessions S3 and S4), we examined how the availability of a nonbinding bonus affects the principals' relative preference for the incentive contract. We implemented the bonus-incentive (BI) treatment for this purpose, where all three contracts could be chosen. Finally, we conducted two sessions (S5 and S6) with control treatments that will be discussed in more detail in Section 4.3.

4. EXPERIMENTAL RESULTS

4.1. *The Trust-Incentive Treatment*

In the TI treatment, principals could choose between a trust contract (w, e^*) and an incentive contract (w, e^*, f). We observed a total of 195 contractual choices. Ten incentive contracts and two trust contracts were rejected, meaning that, in total, the agents made 183 effort choices.

RESULT 1: (a) A clear majority of the contracts in the TI treatment are incentive contracts and the share of incentive contracts increases substantially over time.

(b) The average effort of the agents and the average payoff of the principals are higher under the incentive contracts.

Figure 1 and the following numbers support Result 1(a): 135 (69 percent) of the 195 offered contracts are incentive contracts, while only 60 contracts (31 percent) are trust contracts. However, these numbers fail to demonstrate the strong time trend in the share of incentive contracts shown in Figure 1. While slightly less than 50 percent of the proposed contracts were incentive contracts in the first period of the experiment, this fraction never fell below 70 percent beginning in period 4 and is close to 80 percent of all contracts in the final three periods.[7] Although 71 percent of the principals tried the trust contract at least once, only 33 percent did so in more than three periods. This indicates that most principals experimented somewhat with the trust contract until settling for the incentive contract.

Figure 2 depicts the evolution of average effort levels (and average demanded effort levels) over time for both contract types and illustrates Result 1(b). The figure shows that the average effort is higher in almost all periods

[7]A probit estimation of contractual choices as a function of periods shows that five out of nine period dummies are significant at the 10% level and that the coefficients of the dummies are increasing in seven periods. A random-effects generalized least squares (GLS) regression shows that the probability of choosing an incentive contract goes up by 2.5 percent in each period. This trend is highly significant ($p = 0.005$).

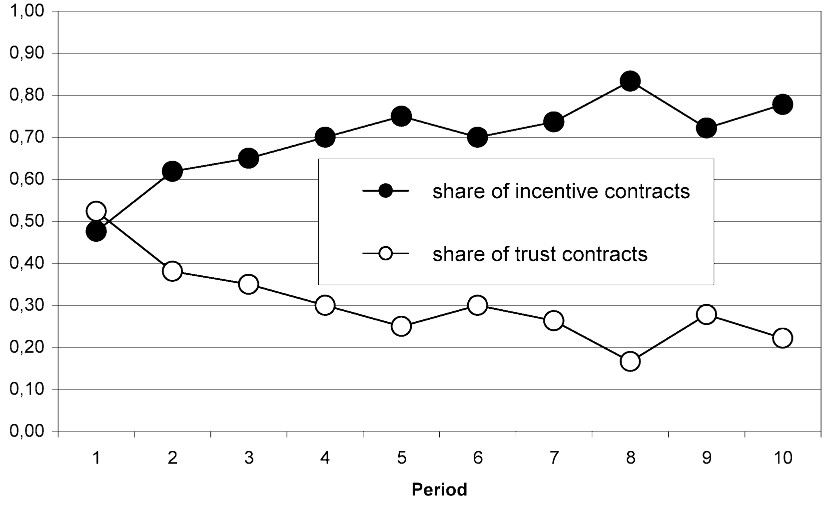

FIGURE 1.—Share of incentive and trust contract (TI treatment).

in the incentive contracts. Moreover, the fraction of trust contracts is already small in those periods in which average effort is somewhat higher in the trust contract, meaning that this is driven by very few observations. The effort differ-

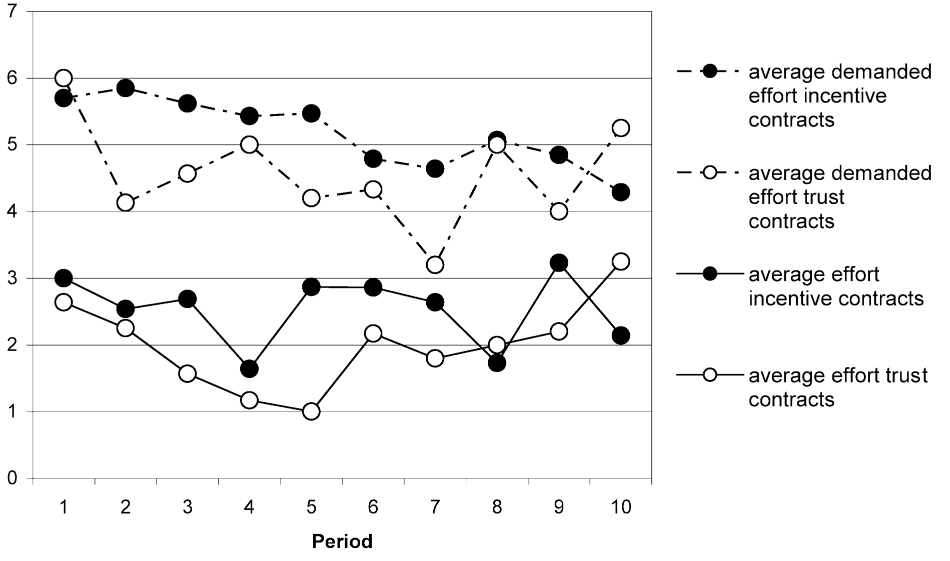

FIGURE 2.—Average effort and average demanded effort in the TI treatment.

ence between ICs and TCs is significant ($p = 0.028$, Mann–Whitney test).[8] This difference in effort levels is also associated with differences in the principals' payoffs. On average, the principals earned a payoff of -0.87 when they proposed an incentive contract and -2.23 when they proposed a trust contract. These payoff differences are, however, not statistically significant ($p > 0.59$, Mann–Whitney test).

Viewed from the perspective of the self-interest model, the rather low profits resulting from the incentive contracts are surprising because—recall from Section 3—the predicted profit is $M^P = 26$. Moreover, it is also surprising that there is such a strong trend toward the incentive contracts in view of the small payoff differences between the incentive and the trust contracts. Why did the principals have such a strong preference for incentive contracts if these contracts performed so poorly? The next result shows that the distinction between incentive-compatible and non-incentive-compatible ICs is crucial in this context.

RESULT 2: (a) Although most incentive contracts stipulate the maximal fine, the majority of incentive contracts violate the no-shirking condition because the principals demand too high effort levels. In the majority of the cases, non-incentive-compatible ICs induce the agents to shirk fully, implying negative payoffs for the principals.

(b) Incentive-compatible ICs are, however, associated with significantly higher payoffs for the principals because the agents shirk much less in these contracts. The large payoff difference between incentive-compatible and non-incentive-compatible ICs is associated with a strong increase in the share of incentive-compatible ICs over time.

Our data support Result 2(a) as follows: The average fine is 12.3, closely approximating the maximal fine of 13. However, the no-shirking condition, $pf \geq c(e^*)$, is violated in 79 (58.5 percent) of the 135 incentive contracts, i.e., principals demanded too high effort levels. Figure 2 also illustrates this fact, showing that the average demanded effort level in the incentive contracts persistently exceeds the maximal enforceable effort $e^* = 4$. We present the agents' effort behavior and the principals' payoffs for incentive-compatible ICs and non-incentive-compatible ICs and TCs in Table III, which shows that non-incentive-compatible ICs are associated with a high rate of shirking and rather low payoffs for the principals. The last row of Table III indicates that there are 79 non-incentive-compatible contracts, that only 1 of these contracts is rejected, and that the agents shirked fully by choosing the minimal effort level $e = 1$ in 48 (62 percent) of the accepted contracts. This high rate of shirking

[8]However, if we look only at the last five periods, the difference in effort levels is no longer significant, which is partly due to the fact that few trust contracts were offered in the last periods.

TABLE III

WAGES, EFFORT, AND PRINCIPALS' PAYOFF IN THE TI TREATMENT

	Incentive-Compatible Incentive Contracts					Non-Incentive-Compatible Incentive Contracts					Trust Contracts				
Wage Offer	No. of Offers	Reject	$e < e^*$	$e \geq e^*$	P's Payoff	No. of Offers	Reject	$e = 1$	$e > 1$	P's Payoff	No. of Offers	Reject	$e = 1$	$e > 1$	P's Payoff
Low $w < 10$	29	8	6	15	8.5	0	n.a.	n.a.	n.a.	n.a.	17	2	15	0	3.7
Medium $10 \leq w \leq 20$	26	1	6	19	9.8	33	1	20	12	−1.4	13	0	9	4	−1.0
High $20 < w$	1	0	0	1	−20.0	46	0	28	18	−12.0	30	0	13	17	−6.4
All	56	9	12	35	8.6	79	1	48	30	−7.6	60	2	37	21	−2.2

has the consequence that the non-incentive-compatible ICs cause on average a loss of -7.6 for the principals.

Figure 2 not only indicates that the average demanded effort in the ICs is too high relative to the enforceable effort level, but also shows that the demanded effort level declines over time. The average demanded effort level in period 1 is close to $e^* = 6$, while it is only slightly above the incentive-compatible level of $e^* = 4$ in the final period. This suggests that the share of incentive-compatible ICs increases over time. In fact, only 10 percent of all ICs are incentive compatible in period one, while this amount exceeds 64 percent of all ICs in period ten. The profit differences between incentive-compatible and non-incentive-compatible ICs provide a natural explanation for this strong time trend. The last row of Table III shows that the average profit in the incentive-compatible ICs is 8.6, which is much larger than the loss of -7.6 in the non-incentive-compatible contracts. Thus, while incentive-compatible ICs are significantly more profitable than trust contracts (Mann–Whitney test, $p = 0.005$), the non-incentive-compatible ICs are less profitable than the trust contracts (Mann–Whitney test, $p = 0.036$). The strong profit differences between the incentive-compatible ICs and the TCs also explain why the share of trust contracts strongly declines over time.

Over time, most principals learned to make the contract incentive compatible, but this was not a trivial task. After all, no principal observed what the other principals did, so everybody had to figure it out on their own. Not all principals managed to do so within the ten rounds of the experiment.

There are two reasons why the incentive-compatible ICs are more profitable than the non-incentive-compatible ICs. First, the principals pay far lower wages when they offer incentive-compatible contracts. Second, although the principals pay less when they offer incentive-compatible ICs, shirking is much less frequent in these contracts. Table III shows that the wage is above $w = 10$ in all 79 ICs that are *not* incentive compatible, while in the majority of the incentive-compatible ICs (29 of 56 cases) the wage is strictly below $w = 10$. This suggests that the principals attempted to elicit reciprocal effort choices from the agents when they proposed non-incentive-compatible contracts. Recall, however, that these attempts frequently failed. This result contrasts sharply with those contracts that meet the no-shirking condition (see the last row of Table III). The agents shirk in only 12 of the 47 accepted incentive contracts.[9]

Table III also indicates that when trust contracts were offered, the principals paid relatively high wages: in 30 of the 60 trust contracts, the principals offered a wage above 20. Thus, the strong decrease in the share of trust contracts and non-incentive-compatible ICs over time caused a decreasing trend in wages over time. The principals offered average wages well above 20 during the first

[9]Note that the number of accepted contracts is given by the sum of the two effort columns. For example, for the incentive-compatible contracts, this sum is given by the 12 contracts with $e < e^*$ plus the 35 contracts with $e \geq e^*$. In all 12 cases with $e < e^*$, the agents chose $e = 1$.

few periods, when the share of incentive compatible ICs was still low. The average wage decreased, however, strongly over time and reached a level of 11.9 in period ten. The strong time trend in the share of incentive-compatible contracts and the average wage suggests that, initially, the principals tried to elicit non-incentive-compatible effort levels by paying generous wages, but as these attempts failed, they converged slowly toward incentive-compatible ICs.

There is a further noteworthy feature in Table III: For the trust contracts the principals' payoff is decreasing in the offered wage. The principals earn 3.7 for wages below $w = 10$, the payoff declines to $M^P = -1.0$ for wages in the middle interval ($10 \leq w \leq 20$), and further diminishes to $M^P = -6.4$ for high wages ($w > 20$). A simple ordinary least squares (OLS) regression of effort on wages also confirms this result, yielding $e = 1.08 + 0.04w + \varepsilon$, where ε denotes the error term (for both coefficients $p < 0.01$ holds). According to this regression, effort significantly increases with wages, but a wage increase of 10 raises effort only by 0.4 and, hence, the expected revenue increases only by 4 units. A similar relationship holds for the non-incentive-compatible ICs, where earnings amount to $M^P = -1.4$ in the middle interval while they correspond to $M^P = -12.0$ for high wages. We summarize this payoff pattern:

RESULT 3: Increasing the generosity of the wage offer as an attempt to induce non-incentive-compatible effort levels decreases the principals' average payoff.

Results 1–3 largely confirm the qualitative predictions of the self-interest model. However, there are also several aspects of the data that are inconsistent with the self-interest model: (i) Principals offer wages that do not hold the agents down to their reservation utilities. On average, a principal offers 35 percent of the surplus that would materialize if the agent chooses the required e^* in the incentive-contract. For incentive compatible ICs, the offered surplus is still 31 percent on average. (ii) The agents reject wages below $w = 10$ in 25 percent of the cases. (iii) The agents shirk by choosing the minimal effort in roughly one-third of the incentive-compatible contracts. All three reasons suggest that fairness concerns do play a role and render the incentive-compatible ICs much less profitable than predicted by the self interest model. Likewise, the fact that the principals initially expressed a strong preference for trust contracts or non-incentive-compatible ICs suggests that they attempted to elicit generous effort choices from the workers. However, the prevailing fairness motives were apparently not strong enough to render these contracts more efficient than the incentive-compatible ICs.

4.2. *The Bonus-Incentive Treatment*

In the BI treatment, principals have an additional contract option: They can announce to pay a (nonbinding) bonus if they are satisfied with the agent's effort. Note that the trust contract is a special type of bonus contract because the principal can always forgo the opportunity to announce a bonus and offer just a fixed wage to the agent. Thus, if a principal wants to propose a trust contract, she sets $b^* = 0$.

We observed a total of 230 contract offers in the BI treatment. Four bonus contracts and two incentive contracts were rejected, leaving 224 accepted contracts. While the incentive contract outperformed the trust contract in the TI treatment, the incentive contract performed very poorly when the principals could choose a bonus contract.

RESULT 4: (a) The overwhelming majority of all contracts in the BI treatment are bonus contracts, while the incentive contract is rarely chosen and the trust contract is never chosen.

(b) The average effort and the average payoff of the principals are much higher in the bonus contract as compared to the incentive contract.

Figure 3 presents the evidence for Result 4(a). Trust contracts do not appear in this figure because they were never chosen. The figure shows the evolution of the share of bonus and incentive contracts over time. Eighty-seven percent of all contracts are already bonus contracts in period one. The share of bonus contracts drops slightly below 80 percent in periods three to five because a few principals experimented with the incentive contract in these periods. However, the share of bonus contracts is roughly 90 percent from period six onward and

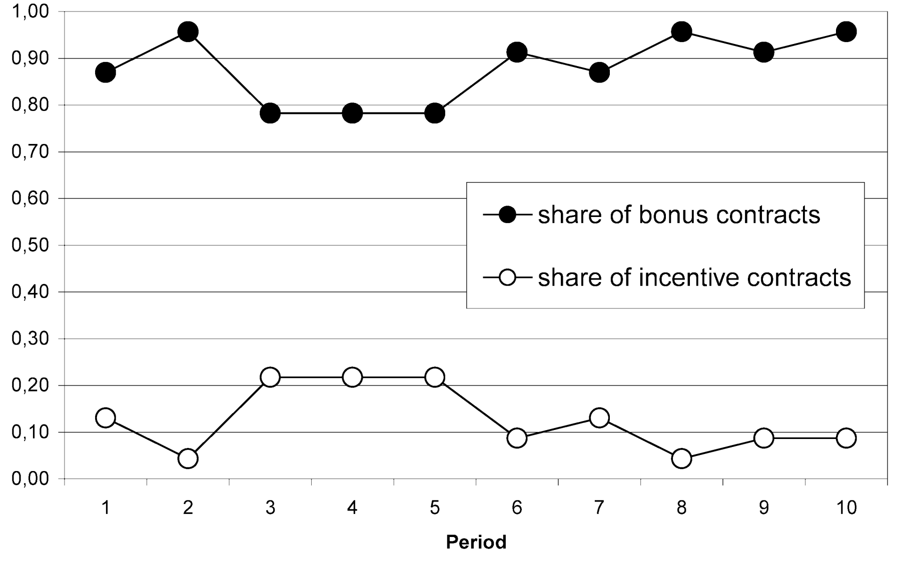

FIGURE 3.—Share of bonus and incentive contracts (BI treatment).

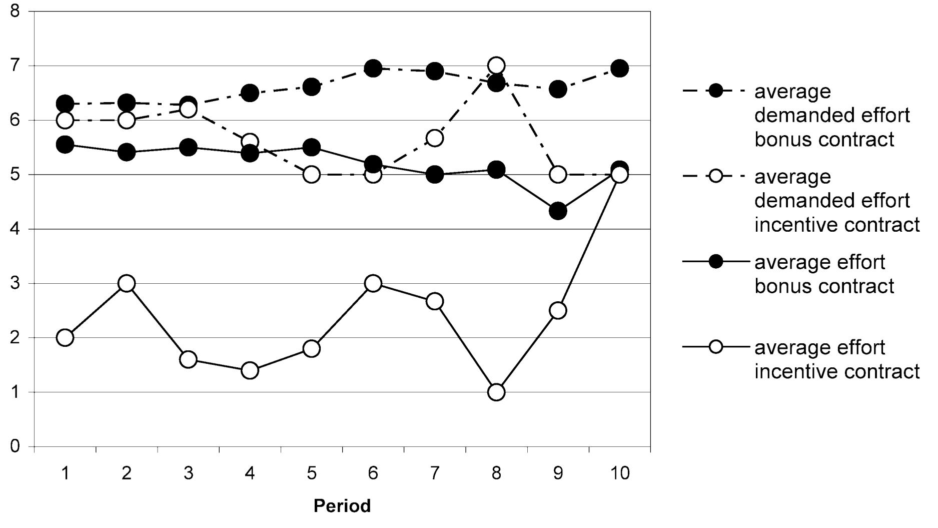

FIGURE 4.—Average effort and average demanded effort in the BI treatment.

even approaches 96 percent in the final period.[10] There can thus be little doubt that principals strongly prefer the bonus contract.

To examine the reasons for this preference, we compare the average effort level in bonus and incentive contracts; see Figure 4. The figure shows that the average effort is considerably higher in the bonus contracts in nine out of ten periods.[11] While the average effort in the incentive contracts is generally between $e = 2$ and $e = 3$, effort in the bonus contracts is, in general, above $e = 5$. This difference across contract types is highly significant ($p < 0.001$, Mann–Whitney test). Figure 4 also indicates that agents' efforts in the bonus contracts are somewhat below the demanded effort level, but the gap between actual and demanded effort levels is much smaller than in the incentive contracts. In fact, as in the TI treatment, many incentive contracts are not incentive compatible. This is indicated by the fact that the demanded average effort always exceeds $e^* = 4$. The large effort differences between the contracts are also translated into large profit differences. Principals' average *profit* from bonus contracts, taken over all ten periods, is 26.8 tokens, while the incentive contract generates an average *loss* of 8.9 tokens. The average profit from bonus contracts is

[10]A probit estimation of contractual choices as a function of periods shows that none of the nine period dummies is significant at the 10 percent level. A random-effects GLS regression shows that the probability of choosing a bonus contract goes up by 1 percent in each period, but this trend is not significant ($p = 0.118$).

[11]The exception is period ten, where the effort difference is negligible. However, there was only one incentive contract in period ten, so that this data point has little relevance for the overall comparison.

always above 20 tokens in each of the ten periods, while the incentive contract causes losses in six of the ten periods. In view of these large profit differences, it is no longer surprising that principals exhibit a strong preference for bonus contracts.

There is an interesting difference in the performance of the incentive contracts across the TI and the BI treatments. In the TI treatment, principals learned to make incentive-compatible contracts over time. In the BI treatment, however, only 5 out of 28 incentive contracts are incentive compatible, and this fraction does not change significantly over time.[12] This explains why the incentive contracts are doing worse in the BI than in the TI treatment. Given that the bonus contract was so attractive (three times as profitable as the incentive-compatible ICs in the TI treatment), it seems that the principals did not bother to learn how to make ICs incentive compatible, but rather switched to bonus contracts.

The higher effort level in the bonus contracts implies a higher surplus. To what extent did the agents receive part of this increase in the surplus relative to the incentive contracts? On average agents earned an income of 14.4 in the incentive contracts, while in the bonus contracts their payoff was 17.8. Thus, on average the agent's payoff is 3.4 tokens higher with a bonus contract than with an incentive contract, while the principal's payoff is 37.1 tokens higher. This shows that the option to pay a bonus yields a substantial efficiency increase *and* causes a sizable change in the distribution of the surplus.

Why does the bonus contract prove to be vastly superior to the incentive contract? Our next result shows that the key to understanding this result is the principals' bonus payments.

RESULT 5: The principals devote a substantial part of the agents' compensation to bonus payments. Moreover, the average bonus increases strongly with respect to the effort level so that nonminimal effort choices are profitable for the agents.

Figures 5 and 6 support Result 5. Figure 5 shows the average wage offered in both the incentive and the bonus contracts; in addition the figure presents the average bonus payments in the BCs. The average wage in the BCs remains in the vicinity of $w = 15$ throughout the whole experiment and the average bonus payment amounts to $b = 10.4$. Thus, on average, the principals pay roughly 40 percent of the agents' total compensation in the form of a bonus. However, this bonus payment strongly depends on the agent's effort (see Figure 6). If the agent provides low effort ($e = 1$ or 2), the average bonus is zero, but the bonus approaches $b = 30$ for high effort levels. The positive slope of the bonus–effort

[12]The five incentive-compatible ICs yielded an average profit of 3.4 to the principals, far less than the average profit of 26.8 of a bonus contract. Thus, even those principals who figured out how to make an IC incentive compatible did much worse than those who used a bonus contract.

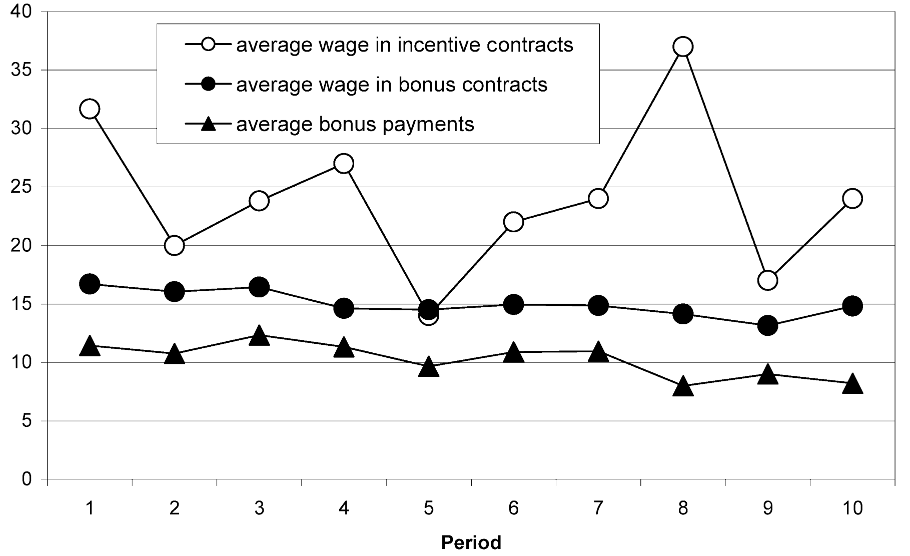

FIGURE 5.—Average wages and average bonus payments over time in the BI treatment.

schedule is also confirmed by the following regressions that relate the bonus payment to the agent's effort e, the demanded effort e^*, the base wage w, and the announced bonus b^* (see Table IV).

Table IV reports the results of several regressions with the associated robust standard errors. The first is a simple OLS regression. In the second regres-

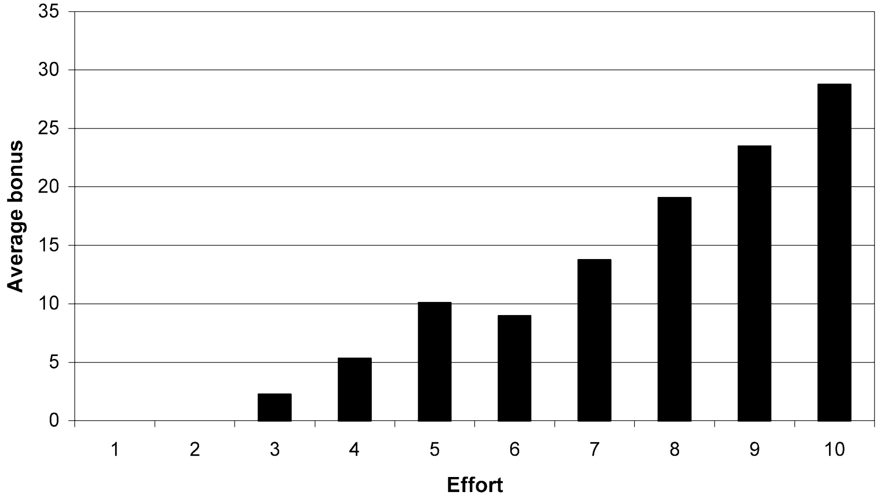

FIGURE 6.—Average bonus as a function of effort in the BI treatment.

TABLE IV

DETERMINANTS OF BONUS PAYMENTS IN THE BI TREATMENT[a]

	Dependent Variable					
	(1)	(2)	(3)	(4)	(5)	(6)
Bonus Payments	Robust Standard Errors	Robust Standard Errors and Clusters	Tobit, Robust Standard Errors, and Clusters	Robust Standard Errors	Robust Standard Errors and Clusters	Tobit, Robust Standard Errors, and Clusters
Constant	−5.58***	−5.58**	−14.55***	−5.58**	−5.58**	−14.55***
	(1.88)	(2.59)	(3.68)	(1.88)	(2.59)	(3.68)
CT				5.19**	5.19	7.35
				(2.56)	(4.30)	(7.23)
Effort	2.86***	2.86***	5.54***	2.86***	2.86***	5.54***
	(0.20)	(0.33)	(0.55)	(0.20)	(0.33)	(0.55)
Effort* *CT*				−0.51*	−0.51	0.04
				(0.29)	(0.56)	(1.03)
Demanded effort	0.33	0.33	−0.59	0.33	0.33	−0.59
	(0.38)	(0.46)	(0.77)	(0.38)	(0.46)	(0.77)
Demanded effort* *CT*				−0.95*	−0.95	−2.03
				(0.54)	(0.64)	(1.32)
Wage	−0.30***	−0.30*	−0.54**	−0.30***	−0.30*	−0.54**
	(0.10)	(0.17)	(0.24)	(0.10)	(0.17)	(0.24)
Wage* *CT*				0.25*	0.25	0.37
				(0.13)	(0.19)	(0.32)
Announced bonus	0.12*	0.12	0.11	0.12**	0.12	0.11
	(0.06)	(0.08)	(0.11)	(0.06)	(0.08)	(0.11)
Announced bonus* *CT*				−0.06	−0.06	−0.04
				(0.07)	(0.09)	(0.18)
Number of observations	198	198	198	376	376	376
Adjusted R^2	0.56	0.56		0.53	0.53	

[a]This table reports the coefficients of OLS and tobit regressions. Robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively. *CT* is a dummy variable for control treatment.

sion we treated the observations of individual principals as separate clusters because they are not independent of each other. Thus, in the second regression the standard errors are based on the assumption that the bonus payments are independent across different principals, but we allow for dependent observations within each cluster. The assumption that the bonus payments are independent across principals is reasonable because a principal could never observe what the other principals did. Moreover, because the bonus payment is the final action in each period, the principals can respond to all previous actions that occurred in the match of that period.

The first two regressions in Table IV show that an increase in the effort level by one unit increases the expected bonus payment by 2.86 tokens. This effect is highly significant in both regressions. Note that 2.86 is higher than the marginal cost of effort for all effort levels $e \leq 7$, i.e., a rational and selfish agent chooses an effort level of $e = 7$ if he faces this bonus–effort relationship. The impact of the demanded effort level is small and not significant, indicating that e^* is considered to be cheap talk. The fixed wage enters the regression with a significantly negative sign, suggesting that if the actual wage increases by 1 token, the principal will reduce the bonus payment by 0.3 tokens on average. The announced bonus enters with a positive, but very small, coefficient that is significant only at the 10 percent level and only in regression (1). An increase in the announced bonus by 10 tokens increases the average actual bonus by only 1.2 tokens. Thus, it seems that principals feel somewhat but not excessively committed to their bonus announcements and that the effort level is the major determinant of the principals' bonus choice. The third column reports a tobit regression that takes into account that the principals could not pay negative bonuses. This has a positive effect on the slope of the bonus–effort relationship, but no effect on the significance of the parameters.

Although the principals respond, on average, quite strongly to increases in the effort level, it is important to notice that there are big differences in individual behavior. This is reported in Table V. In those 162 contracts where the agents chose a nonminimal effort level ($e > 1$), the principals did not pay any bonus at all in 34 cases (21 percent). Among those principals who did pay a bonus, many paid very little even if the agent selected a high effort level. However, there were also many principals who reciprocated high effort levels very generously.

Taken together, the TI and the BI treatments show that the principals strongly prefer the bonus contract. If this contract is not available, they prefer

TABLE V

EFFORT–BONUS RELATIONSHIP IN THE BI TREATMENT (SESSIONS S3 AND S4)[a]

$e \backslash b$	0	1–5	6–10	11–15	16–20	21–25	26–30	31–35	36–40	$\sum$
1	36									36
2	5	1								6
3	9	1	3							13
4	5	6	2	3						16
5	4	3	7	5	3	1				23
6	4	3	10	1	3	1				22
7	6	5	5	6	4	5	5			36
8	0	2	2	6	4	5	4	1		24
9	1	0	2	0	2	3	1	2	2	13
10	0	1	0	0	2	1	1	0	4	9
$\sum$	70	22	31	21	18	16	11	3	6	198

[a]For each possible effort level, the table shows how often a bonus was paid that was either 0 or fell in the given intervals.

the incentive over the trust contract. The same ranking holds in terms of the average effort and the average surplus associated with the three types of contracts. These facts are puzzling from the viewpoint of the self-interest model. Recall that this model captures important qualitative aspects of the TI treatment quite accurately. Although there are several hints in the data that suggest that fairness concerns play a role in the TI treatment, these concerns are apparently too weak to overturn the basic prediction that the principals prefer the incentive contract relative to the trust contract. However, the mere addition of the possibility of announcing and paying a nonbinding bonus—which represents a completely innocuous change from the viewpoint of the self-interest model—suddenly transforms fairness concerns into a powerful determinant of principals' contract choices: Many principals reward generous effort levels with generous bonus payments and thus create powerful incentives for effort provision. In Section 5, we will provide a unified explanation for this puzzle with a fairness model that is based on the assumption of heterogeneous fairness preferences.

4.3. *Control Treatment*

Our experiments were framed as employer–employee relationships, which is natural because many real world employment contracts have similar features to incentive, trust, and bonus contracts. It is also conceivable, however, that this framing affected our subjects' behavior. The work of Samuelson (2001), for example, suggests that the subjects confront decision problems by looking for an analogous situation in the real world and by applying the most suitable real world behavior to the experiment. In the real world, a firm may be much more concerned about its reputation than is a single worker. After all, most firms employ several workers for many periods and it is of crucial importance that they protect their reputation for rewarding good work and honoring implicit contracts. A single worker, on the other hand, has less at stake and may be more inclined to "take the money and run." Of course, reputation effects did not play a role in our experiments because all interactions were one shot and anonymous. However, if the subjects perceived the situation in analogy to a real world employment relationship, this may explain why the trust contracts performed so poorly while the bonus contracts did so well.[13]

To test this hypothesis we did a control experiment with exactly the same structure as the BI treatment, but with a different framing of the instructions. In the control experiment, we had a buyer and seller who want to trade one unit of a good. The quality of the good (and thus the buyer's valuation) depends on the costly effort put in by the seller. To induce the seller to increase the quality of the good, the buyer could offer an incentive contract, a bonus contract, or a trust contract, exactly as in the BI treatment. Note that the above hypothesis

[13]We are grateful to a referee for pointing out this explanation.

predicts the opposite outcome with this framing. A subject who is looking for a real world analogy will identify the seller with a firm that is more concerned about its reputation than a single buyer. Therefore, if frames elicit behaviorally relevant real world analogies, the sellers will apply a reputation heuristic to their quality choices, while the buyers, who have little reason to be concerned with their reputation, will be more likely to take the money and run. If this argument is correct, we should observe that the trust contract, which relies on the sellers' reputation heuristic, does relatively well, while the bonus contract does worse because the buyers have little reason to pay the bonus.

The control treatment was carried out in sessions S5 and S6. Despite the different framing, the experimental results of these sessions are very similar to those of the BI treatment.

RESULT 6: (a) In the control treatment, the overwhelming majority of buyers offer bonus contracts, while the incentive contract is rarely chosen and the trust contract is never chosen. There is no statistically significant difference in the principals' contract choices to the BI treatment.

(b) On average, buyers make significant voluntary bonus payments if the seller's effort is high. There is no statistically significant difference in the bonus–effort relationship between the control and the BI treatment.

(c) As in the BI treatment, bonus contracts elicit significantly higher effort levels and significantly higher average payoffs for both parties than incentive contracts.

Principals chose 179 bonus and 31 incentive contracts in the control treatment, i.e., 85 percent of all contractual offers are bonus contracts. Agents rejected one bonus contract and five incentive contracts. The bonus contract was chosen in 81 percent of all cases in the first three periods; this increased to 87 percent of all cases in the last three periods. A random-effects GLS regression shows that the probability of the bonus contract increases by less than 1 percent per period, but this trend is not significant ($p = 0.249$). This is very similar to the experimental results of the BI treatment. In fact, a Fisher exact probability test shows that the null hypothesis that the contractual offers are drawn from the same distribution as in the BI treatment cannot be rejected ($p = 0.484$).

Result 6(b) is supported by regressions (4) and (5) in Table IV. We pooled the data of the two treatments in these regressions. The dummy variable *CT* is equal to 1 if a data point belongs to the control treatment and 0 if it belongs to the BI treatment. The regressions show that the differences between the two treatments are very small. The effort–bonus and the wage–bonus relationships are a little steeper in the BI treatment than in the control treatment. However, these differences are not significant when the observations of the principals are clustered.

The sellers' average effort level is 4.43 if a bonus contract was offered, more than twice the average effort if an incentive contract was offered (2.16). Thus, as in the BI treatment, the bonus contract is much more efficient. Furthermore, the average payoffs of the buyer and the seller are 1.26 and 12.58, respectively, in the incentive contract, whereas with a bonus contract they earn 22.21 and 16.07, respectively. All of these differences are highly significant (Mann–Whitney tests, $p < 0.015$ in all cases) which confirms Result 6(c).

4.4. *Discussion*

The experiments show that bonus contracts do very well in a one-shot situation in which it is difficult to enforce the agent's effort. There is some anecdotal evidence that contracts for services where quality is difficult to verify or where explicit contracts are not legally enforceable often have the structure of a bonus contract.[14] In a large empirical study on the structure of wage contracts, MacLeod and Parent (1999) showed that bonuses are frequently used and that they dominate the use of piece rates or other explicit incentive pay if jobs are complex and performance is difficult to verify. Scott (2003) offered additional systematic evidence pointing in the same direction. He analyzed a large sample of court cases litigated between 1998 and 2002 in the United States on the grounds of "indefiniteness." He claims that contracts that specify an up-front payment plus an "indefinite" promise of a bonus payment in case of satisfactory performance are quite common in the business world. Sometimes these contracts are litigated because the agent claims that his performance was satisfactory and that he should be paid the bonus, while the principal refuses to pay him. Scott finds that the courts typically refuse to enforce the bonus payment—in particular if the courts have the impression that the contract was deliberately left incomplete. Scott concluded that these bonus contracts rely on reciprocity as an enforcement device, as suggested by our experiments.

5. A THEORETICAL INTERPRETATION

The experimental evidence presented in Section 4 is puzzling at first glance: Why does the incentive contract outperform the trust contract—as standard contract theory predicts—while the bonus contract surpasses the incentive contract, contradicting standard theory? How can the remarkable performance difference between the trust and the bonus contracts be explained, because

[14]For example, when one of us rented a fully furnished house during his sabbatical at Stanford University, his landlord was very concerned how he would treat the antique furniture, water the plants, feed the cat, etc. The contractual solution was a large deposit (in addition to the rent). The understanding was that if the landlord was satisfied with the tenant's behavior, he would repay the deposit; otherwise he would keep it. Note that returning the deposit is like paying a voluntary bonus. It would have been impossible to legally enforce the repayment of the deposit after returning to Germany. Allegations about hiring hit men, drug deals, etc. (as shown in movies) often have the form "half now–half later," again a form of bonus contract. We are grateful to a referee for pointing this out.

after all, both contracts rely on fairness as an enforcement device? How can the poor performance of trust contracts be reconciled with the experimental results of Fehr, Kirchsteiger, and Riedl (1993) that showed that trust contract may do fairly well?

In view of the importance of fairness concerns in our experiments, it is natural to seek an explanation of these puzzles in the context of recently developed fairness models. We show that the fairness approach is consistent with the pattern of contract choices observed in Section 4.[15] In the following discussion, we apply the theory of inequity aversion (Fehr and Schmidt (1999)) to answer the questions raised in the preceding paragraph. We use this theory because it captures important aspects of fairness driven behavior in a tractable way and is consistent with the outcomes of many different experimental games. Even though this theory is very simple and neglects some aspects of the motivational forces of fairness and reciprocity, it gives rise to several interesting new insights and hypotheses for future research.[16]

5.1. *Predicting Contract Choices*

The theory of inequity aversion assumes that some people are very much concerned about inequity, while others care only about their own material payoff. In the two-player case, the utility function of inequity-averse (fair) players is given by

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

$i \in \{1, 2\}$, $i \neq j$, where $x = (x_1, x_2)$ denotes the vector of monetary payoffs, $\beta_i \leq \alpha_i$, $0 \leq \beta_i < 1$, and α_i measures the utility loss that stems from inequality to i's disadvantage, while the term weighted with β_i measures the loss from advantageous inequality. We use a simplified version of this theory by assuming that there are 60 percent self-interested types ($\alpha_i = \beta_i = 0$) and 40 percent fair types. In addition we assume $\alpha_i = 2$ and $\beta_i = 0.6$ for the fair types, i.e., they are prepared to reject contract offers that give them less than 40 percent of the surplus and they are willing to share the surplus equally. This distribution of types is a simplification of the distribution we used in Fehr and Schmidt

[15]However, it is not our aim to explain the dynamic pattern of the data over time. Instead, we focus on the robust behavioral regularities that emerge in all treatment conditions in the final few periods.

[16]It is not our aim here to test the theory of inequity aversion relative to other theories of fairness and reciprocity. Instead we use the theory to acquire a better understanding of the possible mechanisms driving our results. Other theories of fairness and reciprocity (e.g., Bolton and Ockenfels (2000), Charness and Rabin (2002), Cox, Friedman, and Gjerstad (2005), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Levine (1998), or Rabin (1993)) may be able to rationalize the data. However, models of intention-based reciprocity are far more complicated than outcome-based theories like inequity aversion. See Fehr and Schmidt (2003) and Sobel (2005) for surveys of this literature.

(1999).[17] On the basis of these assumptions, our principal agent problem can be analyzed using standard game theoretic tools.[18]

The key reason for the failure of the trust contract in our experiment was that the payment of generous wages was not profitable for the principal. Therefore, only very low effort levels have been ensured under trust contracts, while the use of the fine enabled the principals to enforce higher effort levels in the incentive contracts. This is exactly the result predicted by our fairness model.

PROPOSITION 1—TI Treatment: (a) *Higher wages in the trust contract induce, on average, higher effort levels, but the effort increase is too small to make a wage increase profitable for the principal.*

(b) *In the optimal incentive contract, both the selfish and the fair principals demand the maximal incentive-compatible effort level of* $e^* = 4$ *by imposing the maximal fine* ($f = 13$). *Fair principals distribute the resulting surplus equally by paying a generous wage, while the selfish principals appropriate the whole surplus for themselves by offering the lowest possible wage that a selfish agent accepts. Therefore, the fair principals' optimal contract is always accepted and obeyed, whereas fair agents reject the selfish principals' optimal contract.*

This proposition implies that both types of principals prefer the incentive contracts over the trust contracts because the incentive contract enables them to enforce a higher effort level. To see the intuition for the failure of the trust contract, consider an inequity-averse (i.e., fair) agent who accepts a generous trust contract. He will choose an effort level that equalizes the monetary payoff of the principal with his own monetary payoff:

$$M^P = 10 \cdot e - w = w - c(e) = M^A.$$

Using the implicit function theorem, we get

$$\frac{de}{dw} = \frac{2}{10 + c'(e)}.$$

Thus, e increases with w for an inequity-averse agent; if, however, the fraction of inequity-averse agents in the population is $q = 0.4$, then an increase of w by 1 token increases average effort by at most $\Delta e = 0.4 \cdot (2/11) = 0.07$

[17]Fehr and Schmidt (1999) use four different types: 30 percent of the population are assumed to have $\alpha_i = \beta_i = 0$, 30 percent are assumed to have $\alpha_i = 0.5$ and $\beta_i = 0.3$, 30 percent are assumed to have $\alpha_i = 1$ and $\beta_i = 0.6$, and 10 percent are assumed to have $\alpha_i = 4$ and $\beta_i = 0.6$. It turns out to be very tedious to solve the model for four different types. This is why we simplified the model and use only two different types here. We used the same simplified distribution in Fehr and Schmidt (2004) and Fehr, Kremhelmer, and Schmidt (2005).

[18]The full analysis is somewhat lengthy and therefore is relegated to an appendix that can be found on the *Econometrica Supplementary Materials* web site (Fehr, Klein, and Scmidt (2007a)).

which increases the principal's gross profit by at most $10 \cdot 0.07 = 0.7$ tokens. Hence, a wage increase reduces the principal's expected income, which is what we observed in Table III; higher wages in the trust contract were associated with lower profits. In addition, higher wages generate inequality to the principal's disadvantage whenever a selfish agent chooses $e = 1$. Therefore, both the selfish and the fair principals shy away from trust contracts.

In contrast, the fair principals can enforce an effort level of $e = 4$ without the risk of shirking in an optimal incentive contract. Selfish agents will not shirk because the contract is incentive compatible and fair agents will not shirk because the contract is also fair.[19] The selfish principals also prefer an incentive contract, but they pay a lower wage. Even though this contract is rejected by the fair agents, it enforces an effort level of $e = 4$ and appropriates the whole surplus when facing a selfish agent, and therefore yields a higher expected profit than paying a fair wage that is accepted by both types of the agent.

Thus, the main conclusion from the model of inequity aversion is the same as that from the self-interest model: incentive contracts outperform trust contracts. However, the inequity aversion model is consistent with several other observations in the TI treatment that are not consistent with the self-interest model. First, it explains why low wage offers are frequently rejected. Second, it explains why intermediate but less than fair offers are accepted, with agents subsequently choosing $e = 1$ often, even if the contract is incentive compatible. The fairness model predicts this outcome because fair agents prefer to accept and later shirk on offers that imply a positive but less than fair share of the surplus. Third, the model predicts correctly that incentive contracts are frequently associated with generous wages between 10 and 20 (offered by fair principals). Finally, it offers an explanation as to why many agents choose effort levels larger than 1 in response to generous wage offers in trust contracts and in incentive contracts that are not incentive compatible.

We now turn to the analysis of the bonus contract. The principal has to move twice with a bonus contract: first when offering the contract and second when deciding on the size of the bonus to be paid. Thus, the bonus contract induces a signaling game in which the agent may take the contract offer as a signal about the principal's type.

The salient facts in the BI treatment are that the principals preferred the BC almost universally, that the average effort is higher in the BC than in the IC, and that many principals paid a generous bonus in response to high effort lev-

[19]The induction of effort levels above $e = 4$ is not profitable in incentive contracts because the percentage of fair agents is too small to render this profitable. Note that for $e > 4$, $c'(e) \geq 2$. Hence, the marginal revenue of a unit increase in wages at $e = 4$ equals $10 \cdot q \cdot (de/dw) = 10 \cdot q \cdot (2/12)$, which exceeds 1 for $q > 12/20 = 6/10$. An effort level $e \leq 4$ can be implemented with an incentive contract at a lower risk of suffering from inequality to the principal. Note also that even if $q > 0.6$, the inequity-averse principals need not offer generous wages because they may still be afraid to suffer from the inequality caused by the selfish agents.

els. All these facts are implied by the following proposition derived from our model.

PROPOSITION 2—Bonus Contracts: (a) *No separating equilibrium exists in which the selfish principals' contract offer differs from that of the fair principals' offer.*

(b) *If a wage increase in a bonus contract is not interpreted as a signal that the principal is more likely to be selfish, there is a unique pooling equilibrium outcome in which both types of principals offer* $w = 15$. *The selfish agent chooses* $e = 7$ *and is rewarded by the fair principal with a bonus of* 25. *The fair agent chooses* $e = 3$ *and is rewarded by the fair principal with a bonus of* 1. *The selfish principal never pays a bonus. Thus, the expected bonus is* 6.16 *and the expected effort level is* 5.4.

This proposition implies that both types of principals prefer the BC over the IC. In the unique pooling equilibrium, the average effort is given by $(1-q)\cdot 7 + q\cdot 3 = 0.6\cdot 7 + 0.4\cdot 3 = 5.4$, which is higher than the effort in the optimal incentive contract. Because the fair principals share the surplus of a contract equally regardless of the type of contract and because the bonus contract generates a higher surplus, the fair principals unambiguously prefer the bonus contract. Moreover, because the selfish principals reap the benefits of the high effort level in a bonus contract without actually paying the bonus, they also prefer the bonus contract. Thus, the selfish principals can exploit the fact that there are fair principals, because the latter pay a generous bonus, which provides strong performance incentives for the selfish agents.

To provide an intuition for the absence of a separating equilibrium, suppose that the fair and the selfish principal offer different contracts. In this case, the agents know from the contract offer whether they face a fair principal or a selfish principal. If they face a fair principal, they choose $e = 10$ because this principal pays a bonus that distributes the surplus equally. If they face a selfish principal, they choose $e = 1$. Hence, a selfish principal always wants to mimic the contractual offer of the fair principal.

To derive the agents' effort choice, consider the last stage of a bonus contract. It is obvious that a selfish principal will not pay a bonus while a fair principal pays a bonus that equalizes payoffs:

$$10\cdot e - w - b = w + b - c(e).$$

Using the implicit function theorem, we get

$$\frac{db}{de} = \frac{10 + c'(e)}{2}.$$

In the pooling equilibrium the expected monetary payoff of the agent as a function of e is given by

$$M^A(e) = q\cdot[w + b(e) - c(e)] + (1-q)\cdot[w - c(e)].$$

Differentiating with respect to e yields

$$\frac{dM^A}{de} = q \cdot \frac{db}{de} - c'(e) = q \cdot \frac{10 + c'(e)}{2} - c'(e).$$

This expression is positive if q is large enough compared to $c'(e)$. Recall that, according to the cost schedule in Table I, $1 \leq c'(e) \leq 4$. The critical value for q that ensures $dM^A/de \geq 0$ is $2/11 \approx 0.18$ for $c' = 1$; it is 0.33 for $c' = 2$, 0.46, for $c' = 3$, and 0.57 for $c' = 4$. Hence, in a pooling equilibrium, where the agents believe that they face a fair principal with probability $q = 0.4$, selfish agents will choose the maximal effort level for which the marginal effort cost does not exceed 2, that is, they choose $e = 7$.

It is important to note that the theory implies that only self-interested agents choose $e = 7$. Fair agents choose a lower effort because they suffer from inequity if the bonus is not paid (the fair agent chooses $e = 3$ if $w = 15$).[20] Thus, the presence of fair principals induces selfish agents to choose high effort levels, while the presence of selfish principals induces the fair agents to provide low effort levels. This is an interesting example of the sometimes surprising effects that arise in a heterogeneous population with fair and selfish subjects.

Proposition 2 not only rationalizes the contract choices in the BI treatment, but also offers relatively precise quantitative predictions of the average values of wages, bonuses, and effort. The average wage offered with a bonus contract is 15.0, the average bonus is 10.4, and the average actual effort level is 5.2, all fairly close to the predictions of Proposition 2. However, it also has to be said that there is a lot of noise in the data and that the model does less well in explaining individual behavior. There are some principals whose behavior is consistent with the model and who either choose $b = 0$ in all periods or a bonus that (roughly) equalizes payoffs. However, there are also principals who pay a positive but smaller bonus and many of them do not behave consistently over time.

5.2. *Is a Bonus Contract Always Better than a Trust Contract?*

Under the assumptions about the value and cost functions that we implemented in our experiment, the model of inequity aversion predicts that a bonus contract is better than an incentive contract, which in turn outperforms a trust contract. How general is this ranking?

Clearly, the relative performance of the incentive contract depends on the contractual environment. If it is possible to enforce an effort level that is sufficiently close to e^{FB} and if contracting costs are not too large, then the incentive

[20]The agent has to trust the principal in a bonus contract. Bohnet and Zeckhauser (2004) have conducted interesting experiments that indicate that social preferences inhibit trusting behavior. They provided evidence that both inequity aversion and the aversion against being cheated tend to reduce trusting behavior.

contract is optimal and should be used by both the self-interested and the fair principal. On the other hand, if the enforcement technology is sufficiently imperfect, an implicit bonus contract will outperform the incentive contract, as we have shown in Section 4.

What can be said about the comparison between bonus and trust contracts? In a trivial sense, the bonus contract must always be at least as good as the trust contract. After all, the bonus contract can simply mimic the trust contract by offering a high wage up front and setting $b^* = 0$. However, subjects did not do this. Almost all principals who offered a bonus contract paid a modest wage up front and used the announcement of the bonus as an incentive device. This induced much higher effort levels than paying a generous wage up front in a trust contract.

It this section we will show that the superiority of the bonus contract over the trust contract is a general result that is independent of the specific parameters and functional forms that we employed in the experiments. In fact, we can show that under the general assumptions on $v(e)$ and $c(e)$ of Section 2, a contract that relies on the promise of a voluntary bonus payment *always* induces a (weakly) higher effort level than a contract that pays a generous wage up front[21]:

PROPOSITION 3: *For all $q \in (0, 1)$, the expected effort chosen by an agent under an optimal bonus contract is (weakly) higher than the expected effort induced by a principal in an optimal trust contract. It is strictly higher if the average effort induced by the optimal bonus contract is strictly larger than the minimum effort level.*

The intuitive reason for the superiority of the bonus contract is that the expected cost of trusting is much lower with a bonus than with a trust contract. The principal in a trust contract has to pay a generous wage up front. To induce a fair agent to provide a nonminimal effort level of e, the principal has to offer a wage that covers the agent's effort cost $c(e)$ plus half of the surplus that is generated if the agent chooses e. Thus, the principal makes a risky wage "investment" of $w(e) = c(e) + [v(e) - c(e)]/2$, which yields only an effort return if the principal faces a fair agent. In contrast, the agent in a bonus contract has to trust first by providing effort so that the agent's risky "investment" is given by $c(e)$. Because $c(e) < v(e)$ implies $c(e) < w(e)$ the risk of trusting is always lower for the agent in a bonus contract than for the principal in a trust contract. As a consequence, the expected marginal payoff of a higher effort level is also larger for an agent in a bonus contract than for a principal in a trust contract, which implies that the effort level the agent chooses in a bonus contract is higher than the effort level that the principal induces in a trust contract.

[21]There we assume $v' > 0$, $c' > 0$, $v'' < 0$, $c'' \geq 0$, and $e^{FB} \geq \underline{e}$. The proof of Proposition 3 can be found on the *Econometrica Supplementary Materials* web site (Fehr, Klein, and Schmidt (2007a)).

Proposition 3 suggests, therefore, a more general principle: If a relationship is based on trust in the other player's fairness, the risk of trusting should be borne by the party for whom the cost of trusting is lower.

5.3. *Comparison to Fehr, Kirchsteiger, and Riedl*

Fehr, Kirchsteiger, and Riedl (1993) introduced a gift exchange game in which principals were restricted to offer trust contracts. In their experiments, as in ours, there was a significant positive relationship between wages the principals offered and effort the agents spent. In contrast to our results, however, the wage-effort relationship was sufficiently steep to make it profitable for the principals to offer high wages, and many principals did so. The FKR (1993) results have been replicated by several other studies that were also based on the payoff function used by FKR (Charness (2004), Charness, Frechette, and Kagel (2004), Fehr, Kirchsteiger, and Riedl (1998), Fehr, Kirchler, Weichbold, and Gächter (1998), Hannan, Kagel, and Moser (2002)).[22]

The main difference between FKR (1993) and our paper is the principal's payoff function: FKR used the payoff function $M^P = (v - w) \cdot e = (126 - w) \cdot e$, while we used $M^P = v \cdot e - w = 10 \cdot e - w$. In addition, the agent had to choose an effort level $e \in \{0.1, 0.2, \dots, 1\}$, whereas $e \in \{1, 2, \dots, 10\}$ in our experiment. If we apply our model (with the same assumptions as in Section 5.1) to the setting of FKR, we find that the model is consistent with FKR (see our web based appendix). In particular, the model predicts the wage–effort schedule in FKR (1993, Table 2) very well. The predicted schedule implies that the principals in FKR can increase their monetary payoff by paying generous wages. Intuitively, with the payoff function $M^P = (v - w) \cdot e$, the payment of high wages is less costly for the principal if the agent chooses a low effort. This makes high wages less risky and, therefore, more profitable for the principal than in our experiment. Thus, from the perspective of our model, there is no contradiction between the experimental results reported in this paper and the results reported by FKR and others. At a more quantitative level, the principals in FKR pay, on average, a wage of 72, which is associated with an average effort of 0.45. Our model predicts an average wage of 63 and an average effort of 0.46.

6. CONCLUSIONS

Our experiments have shown that fairness concerns may have important consequences for the optimal provision of incentives. Incentive contracts that are

[22]The paper by Charness, Frechette, and Kagel (2004) also finds an interesting presentation effect. If agents receive payoff information in the form of a comprehensive payoff table, which provides information about the agent's and the principal's payoff for all feasible wage–effort combinations, reciprocal effort behavior is reduced significantly. Results like these remind us that the psychology of incentives and social preferences is sufficiently complex that all prevailing models fail to provide a complete account of all the subtle facts.

optimal when there are only selfish actors perform less well when some agents are concerned about fairness. On the other hand, implicit bonus contracts that cannot work when all actors are selfish provide powerful incentives and become superior when there are also fair-minded players. Our results indicate that the principals understand that fairness matters and predominantly choose the superior bonus contract that relies on fairness as an enforcement device.

There are several other points that deserve to be emphasized. First, the principals converge toward the most efficient contract in the set of available contracts. This observation is important because the "efficiency principle" provides the basis for much of modern contract theory. However, it remains to be seen whether this observation extends to other—more complicated—environments. Second, it is important to remember that only some subjects are concerned about fairness. A considerable percentage of subjects also seem to be mainly interested in their own material payoff. Whether fairness motives provide a good enforcement device depends on the percentage of fair persons in the population *and* on the strategic situation in which the subjects interact. We have shown that fairness concerns are too weak for contract enforcement in a setting where the trust contract competes with the incentive contract, but they are strong enough if the bonus contract becomes available. This asymmetry in the impact of fairness concerns is due to the fact that it is less costly for the agent to trust in the bonus contract than for the principal to trust in the trust contract. Third, our theoretical results show that simple and tractable models of fairness can yield interesting and nonobvious insights into the problems of contract choice and incentive provision. Our fairness model is consistent with the contract choice patterns in the data. In addition, it provides relatively accurate quantitative predictions of the average values of wages, bonuses, and effort of the bonus contract.

Finally, our experiments and the theoretical analysis show that the presence of fair types does not automatically provide a solution to every contracting problem and may sometimes even exacerbate incentive and contracting problems. Fair types are much more afraid of being taken advantage of. For this reason, fairness preferences inhibit trusting behavior because trust typically involves a risk of being cheated. A recent paper by Bohnet and Zeckhauser (2004) provides clean evidence in favor of this view. Our theoretical analysis also shows that fair agents respond less to the implicit incentives provided by a bonus contract than selfish agents, because fair agents experience *additional* disutility if the bonus is not paid, while selfish agents "only" suffer from the reduced material payoff. This finding shows that the presence of fair players may complicate the task of incentive provision because—in addition to the conventional incentive compatibility constraints—the "fairness compatibility" of the contract also has to be taken into account.

Our experiments show that concerns for fairness have an important impact on the actual and the optimal design of contracts. Traditional contract theory has neglected these effects, but they have to be taken into account if we want to fully understand the functioning of real world contracts and the associated incentive schemes. Our theoretical analysis shows that it is possible to model these effects explicitly. The model of inequity aversion that we used is clearly very simple and ignores certain aspects of human behavior. Although it may not offer a good description of the behavior of individual subjects, the contract choice predictions are rather accurate. On average, subjects behave *as if* they were motivated by inequity aversion. The model not only helps to organize and interpret the data, it also give rise to interesting and important new hypotheses that can be tested experimentally. This approach is a first step to developing richer models that may become part of "behavioral contract theory."

Institute for Empirical Research in Economics, University of Zurich, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland, and Collegium Helveticum, CH-8092 Zürich, Switzerland; efehr@iew.unizh.ch,

Dept. of Economics, University of Munich, Ludwigstrasse 28, D-80539 Muenchen, Germany,

and

Dept. of Economics, University of Munich, Ludwigstrasse 28, D-80539 Muenchen, Germany; klaus.schmidt@Lrz.uni-muenchen.de.

Manuscript received April, 2004; final revision received July, 2006.

REFERENCES

ANDERHUB, V., S. GÄCHTER, AND M. KÖNIGSTEIN (2002): "Efficient Contracting and Fair Play in a Simple Principal–Agent Experiment," *Experimental Economics*, 5, 5–27. [124]

BERG, J., J. W. DICKHAUT, AND K. A. MCCABE (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122–142. [124]

BOHNET, I., AND R. ZECKHAUSER (2004): "Trust, Risk and Betrayal," *Journal of Economic Behavior and Organization*, 55, 467–484. [148,151]

BOLTON, G. E., AND A. OCKENFELS (2000): "A Theory of Equity, Reciprocity and Competition," *American Economic Review*, 100, 166–193. [144]

CAMERER, C. F., AND K. WEIGELT (1988): "Experimental Tests of a Sequential Equilibrium Reputation Model," *Econometrica*, 56, 1–36. [124]

CHARNESS, G. (2004): "Attribution and Reciprocity in an Experimental Labor Market," *Journal of Labor Economics*, 22, 665–688. [124,150]

CHARNESS, G., AND M. RABIN (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817–869. [144]

CHARNESS, G., G. FRECHETTE, AND J. KAGEL (2004): "How Robust Is Laboratory Gift Exchange?" *Experimental Economics*, 7, 189–205. [150]

CHAUDHURI, A. (1998): "The Ratchet Principle in a Principal–Agent Game with Unknown Costs: An Experimental Analysis," *Journal of Economic Behavior and Organization*, 37, 291–304. [124]

CHIAPPORI, P. A., AND B. SALANIÉ (2003): "Testing Contract Theory: A Survey of Some Recent Work," in *Advances in Economic Theory*, *Eighth World Congress of the Econometric Society*, Vol. 1, ed. by M. Dewatripont, L. P. Hansen, and S. Turnovski. Cambridge, U.K.: Cambridge University Press, 115–149. [123,124]

COOPER, D. J., J. H. KAGEL, W. LO, AND Q. L. GU (1999): "Gaming Against Managers in Incentive Systems: Experiments with Chinese Students and Chinese Managers," *American Economic Review*, 89, 781–804. [124]

COX, J., D. FRIEDMAN, AND S. GJERSTAD (2005): "A Tractable Model of Reciprocity and Fairness," Mimeo, University of Arizona at Tucson. [144]

DEJONG, D. V., R. FORSYTHE, R. J. LUNDHOLM, AND W. C. UECKER (1985): "A Laboratory Investigation of the Moral Hazard Problem in an Agency Relationship," *Journal of Accounting Research*, 23, 81–120. [124]

DUFWENBERG, M., AND G. KIRCHSTEIGER (2004): "A Theory of Sequential Reciprocity," *Games and Economic Behavior*, 47, 268–298. [144]

FALK, A., AND U. FISCHBACHER (2006): "A Theory of Reciprocity," *Games and Economic Behavior*, 54, 293–315. [144]

FEHR, E., S. GÄCHTER, AND G. KIRCHSTEIGER (1997): "Reciprocity as a Contract Enforcement Device," *Econometrica*, 65, 833–860. [124]

FEHR, E., G. KIRCHSTEIGER, AND A. RIEDL (1993): "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics*, 58, 437–460. [123,124,144, 150]

——— (1998): "Gift Exchange and Reciprocity in Competitive Experimental Markets," *European Economic Review*, 42, 1–34. [150]

FEHR, E., G. KIRCHLER, A. WEICHBOLD, AND S. GÄCHTER (1998): "When Social Norms Overpower Competition—Gift Exchange in Experimental Labor Markets," *Journal of Labor Economics*, 16, 324–351. [150]

FEHR, E., A. KLEIN, AND K. M. SCHMIDT (2007a): "Supplement to 'Fairness and Contract Design': Appendix," *Econometrica Supplementary Material*, 75, http://www.econometricsociety.org/ecta/supmat/5182proofs.pdf. [125,145,149]

——— (2007b): "Supplement to 'Fairness and Contract Design': Experimental Data," *Econometrica Supplementary Material*, 75, http://www.econometricsociety.org/ecta/supmat/5182data.pdf. [125,128]

——— (2007c): "Supplement to 'Fairness and Contract Design': Instructions," *Econometrica Supplementary Material*, 75, http://www.econometricsociety.org/ecta/supmat/5182instructions.pdf. [128]

FEHR, E., AND K. M. SCHMIDT (1999): "A Theory of Fairness, Competition and Co-Operation," *Quarterly Journal of Economics*, 114, 817–868. [123,144,145]

——— (2003): "Theories of Fairness and Reciprocity: Evidence and Economic Applications," in *Advances in Economic Theory*, *Eighth World Congress of the Econometric Society*, Vol. 1, ed. by M. Dewatripont, L. P. Hansen, and S. Turnovski. Cambridge, U.K.: Cambridge University Press, 208–257. [144]

GÜTH, W., W. KLOSE, M. KÖNIGSTEIN, AND J. SCHWALBACH (1998): "An Experimental Study of a Dynamic Principal–Agent Relationship," *Managerial and Decision Economics*, 19, 327–341. [124]

GÜTH, W., M. KÖNIGSTEIN, J. KOVACS, AND E. ZALA-MEZO (2001): "Fairness Within Firms: The Case of One Principal and Multiple Agents," *Schmalenbach Business Review*, 53, 82–101. [124]

HANNAN, L., J. KAGEL, AND D. MOSER (2002): "Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences and Effort Requests on Behavior," *Journal of Labor Economics*, 20, 923–951. [124,150]

KESER, C., AND M. WILLINGER (2000): "Principals' Principles when Agents' Actions Are Hidden," *International Journal of Industrial Organization*, 18, 163–185. [124]

LEVINE, D. (1998): "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1, 593–622. [144]

MACLEOD, W. B., AND PARENT, D. (1999): "Job Characteristics and the Form of Compensation," in *Research in Labor Economics*, Vol. 18, ed. by S. W. Polachek. London: JAI Press, 177–242. [143]

PRENDERGAST, C. (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, 7–63. [123]

RABIN, M. (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83, 1281–1302. [144]

SAMUELSON, L. (2001): "Analogies, Adaptation and Anomalies," *Journal of Economic Theory*, 97, 320–366. [141]

SCHOTTER, A., AND H. NALBANTIAN (1997): "Productivity under Group Incentives: An Experimental Study," *American Economic Review*, 87, 314–340. [124]

SCHOTTER, A., C. BULL, AND K. WEIGELT (1987): "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95, 1–33. [124]

SCOTT, R. E. (2003): "A Theory of Self-Enforcing, Indefinite Agreements," *Columbia Law Review*, 103, 1641–1699. [143]

SOBEL, J. (2005): "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43, 392–436. [144]