



LUDWIG-  
MAXIMILIANS-  
UNIVERSITÄT  
MÜNCHEN

VOLKSWIRTSCHAFTLICHE FAKULTÄT



Fehr, Ernst; Klein, Alexander und Schmidt, Klaus M.:  
Fairness, Incentives and Contractual Incompleteness

Munich Discussion Paper No. 2001-7

Department of Economics  
University of Munich

Volkswirtschaftliche Fakultät  
Ludwig-Maximilians-Universität München

Online at <https://doi.org/10.5282/ubm/epub.18>

# Fairness, Incentives and Contractual Incompleteness\*

Ernst Fehr<sup>a)</sup>

University of Zurich, CESifo and CEPR

Alexander Klein<sup>b)</sup>

University of Munich

Klaus M. Schmidt<sup>c)</sup>

University of Munich, CESifo and CEPR

This version: February 2001

**Abstract:** We show that concerns for fairness may have dramatic consequences for the optimal provision of incentives in a moral hazard context. Incentive contracts that are optimal when there are only selfish actors become inferior when some agents are concerned about fairness. Conversely, contracts that are doomed to fail when there are only selfish actors provide powerful incentives and become superior when there are also fair-minded players. These predictions are strongly supported by the results of a series of experiments. Furthermore, our results suggest that the existence of fair actors may be an important reason why many contracts are left deliberately incomplete.

**Keywords:** Incentive Contracts, Moral Hazard, Fairness, Reciprocity, Incomplete Contracts

**JEL Classification Numbers:** C7, C9, J3

---

\* We would like to thank Colin Camerer, Dan Friedman, Paul Milgrom, Jean Tirole and seminar participants at Bocconi University, the California Institute of Technology, the University of California at Berkeley and at Santa Cruz, the University of Munich, Stanford University, the University of Toulouse, the University of Washington and University College London for helpful comments and suggestions. Susanne Kremhelmer 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 grant SCHM-1196/4-1 is gratefully acknowledged. Ernst Fehr also gratefully acknowledges support from the Swiss National Science Foundation (project number 1214-05100.97) and the Network on the Evolution of Preferences and Social Norms of the MacArthur Foundation and the EU-TMR Research Network ENDEAR (FMRX-CTP98-0238).

<sup>a)</sup> Ernst Fehr, Institute for Empirical Research in Economics, University of Zurich, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland, email: [efehr@iew.unizh.ch](mailto:efehr@iew.unizh.ch).

<sup>b)</sup> Alexander Klein, Department of Economics, University of Munich, Ludwigstrasse 28, D-80539 Muenchen, Germany, email: [alexander.klein@Lrz.uni-muenchen.de](mailto:alexander.klein@Lrz.uni-muenchen.de).

<sup>c)</sup> Klaus M. Schmidt, Department of Economics, University of Munich, Ludwigstrasse 28, D-80539 Muenchen, Germany, email: [klaus.schmidt@Lrz.uni-muenchen.de](mailto:klaus.schmidt@Lrz.uni-muenchen.de) (corresponding author).

## I. Introduction

This paper shows how concerns for fairness affect the optimal provision of incentives. Traditional contract theory is based on the assumption that principals and agents are solely interested in their own material payoffs. While this assumption may be an accurate description of the behavior of many people, it is clearly contradicted if applied to *all* people. In fact, there is by now considerable evidence indicating that a substantial fraction of the people also cares about reciprocal fairness (see e.g. the surveys of Camerer 2000, and Fehr and Schmidt 2000).<sup>1</sup> In this paper we show theoretically and by a series of experiments that the existence of a fraction of fair-minded people has dramatic consequences for the optimality of different types of contracts. Based on the recently developed theory of inequity aversion (Fehr and Schmidt, 1999) we show theoretically that contracts that are optimal when all actors are selfish may be much less efficient when there is a minority of people who care about fairness. Furthermore, contracts that would be very inefficient if all actors were selfish, may achieve astonishingly high levels of efficiency when there are some fair-minded people. Thus, our theoretical results strongly suggest that concerns for fairness should be taken into account in the design of optimal incentive schemes.

Our experiments provide substantial support for these theoretical predictions. The type of contract that is optimal according to standard contract theory turns out to be far less efficient than predicted by this theory, while contracts that are predicted to be very inefficient if all agents are purely self-interested turn out to be superior. The experimental results clearly suggest that this reversal in contract efficiency is due to the existence of fair subjects. The principals in our experiments seem to understand this quite well. A large majority of them chooses the contract that is predicted to be optimal in the presence of a minority of fair agents. Those who choose the contract predicted by the standard theory do very poorly.

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 spend effort which is personally costly to the agent. Effort is observable by both parties. However, it can be verified to the courts only if the principal invests in a verification technology. If she makes this investment, she can offer an “incentive contract” to the agent, which fines the agent for unsatisfactory performance. The problem

---

<sup>1</sup> Reciprocally fair people respond to friendly actions in a friendly manner, but they retaliate in response to hostile actions. These friendly and hostile responses also prevail in one-shot situations and when they are costly for the responder (see Rabin 1993 for this definition of reciprocal fairness).

with the incentive contract is that the verification technology is imperfect and that the fine that can be imposed on the agent is bounded, so the highest implementable effort level is positive but falls short of the efficient effort level. Alternatively, the principal can offer a “bonus contract” which does not rely on effort verification and enforcement by third parties. Instead, the principal announces to make a voluntary bonus payment if the agent’s effort is to her satisfaction. The problem with the bonus contract is that the principal cannot be forced to pay the bonus. Note that the bonus contract is highly incomplete because it does neither constrain the principal’s bonus payment nor the agent’s effort level. The incentive contract, in contrast, is less incomplete because it makes pay explicitly contingent on effort.

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 minimal effort levels. Thus, according to standard contract theory, the bonus contract is doomed to fail, while the incentive contract is predicted to do much better. The incentive contract does not achieve first best efficiency, but at least it induces the agents to choose a medium level of effort. Yet, we show theoretically that if there is a minority of reciprocally fair subjects then the incentive contract performs worse than predicted by standard contract theory, while powerful incentives can be given by the incomplete bonus contract. In fact, in our experiments the bonus contract turned out to be much more efficient and much more profitable than the incentive contract.

The experiment can also be used to illustrate some of the intricate and often surprising effects that arise when some (but not all) people are fair-minded. A fair principal will pay the bonus if the agent worked sufficiently hard. Thus, if the fraction of fair principals is not too small, it is profitable for a purely self-interested agent to choose a high effort level. This makes it profitable for a selfish principal to mimic the contract offered by the fair principals in order to benefit from the high effort levels of the selfish agents under a bonus contract without paying, in fact, the bonus. Yet, fair-minded agents strongly dislike being the sucker, i.e., they experience *additional* disutility if they work hard but do not receive the bonus. If the principal cannot credibly signal that she is going to pay the bonus, fair agents are *not* willing to work under a bonus contract. Hence, it may happen that the existence of fair principals induces the selfish agents to perform well while the existence of selfish principals induces the fair agents to provide little effort under a bonus contract. Thus, surprisingly, too many fair agents can be detrimental to the efficiency of a bonus contract.

Our paper can also be seen as a contribution to the literature on incomplete contracts. The simplicity and incompleteness of many contracts is a puzzle for traditional contract theory which predicts that contracts should depend on all verifiable information that contains statistical information about the agent's action or type (see, e.g., Holmström, 1982, Laffont and Tirole, 1993). However, actual contracts frequently specify important obligations of the contracting parties in fairly vague terms, and they do not tie the parties' monetary payoffs to measures of performance that would be available at a relatively small cost.<sup>2</sup> Thus, it seems that many contracts are left deliberately incomplete.

There have been many attempts to explain contractual incompleteness. Some papers, e.g. Aghion, Dewatripont, and Rey (1994), Nöldeke and Schmidt (1995), or Edlin and Reichelstein (1996) argue that in some situations a properly designed incomplete contract that is completed by renegotiation implements the first best already, so there is no need to write a more complete contract. Other papers (e.g. Che and Hausch, 1998, Segal, 1999, and Hart and Moore, 1999) show that, although the best incomplete contract does not implement the first best, a more complete contract cannot improve the situation and is of no value to the contracting parties. Spier (1992) and Allen and Gale (1992) argue that a principal may offer an incomplete contract, because a more complete contract would be interpreted as an unfavorable signal by the other party.

Our paper differs from this literature by pointing out that an incomplete contract may rely on fairness and reciprocity as an enforcement device. Moreover, the experimental results show that this enforcement mechanism can be very effective, and that the large majority of principals used it quite successfully. The incomplete bonus contract is more efficient than our less incomplete incentive contract *because* it is less complete and thus gives more freedom to the parties to reciprocate.

This last insight is closely related to the literature on implicit incentives in repeated relationships. MacLeod and Malcomson (1989) consider a principal agent game similar to the one considered in this paper where the agent is free which effort to spend while the principal is free what bonus to pay. If this game is infinitely repeated, then there exists a subgame perfect equilibrium in which the agent works efficiently and the principal pays a generous bonus in every period. The equilibrium is sustained by the threat that, if any party deviates, the inefficient equilibrium where the agent does not work and the principal does not pay a bonus will be played in all subsequent periods.

---

<sup>2</sup> For example, a typical contract for a university professor does not make the salary directly contingent on easily measurable and verifiable measures of performance such as citations, teaching ratings or the placement of Ph.D. students.

Baker, Gibbons and Murphy (1994), Schmidt and Schnitzer (1995) and Bernheim and Whinston (1998) extend this analysis and show in different setups that it may be beneficial not to contract on some variables that are easily verifiable and contractible in order to give more discretion to the parties to retaliate to a deviation from the implicit contract. However, all of these models are based on repeated game effects. In contrast, our paper shows that implicit contracts can be sustained even in a one-shot principal-agent game if some players are concerned about fairness and reciprocity.

Finally, our theoretical and experimental results provide strong support for Holmström and Milgrom's (1991) hypothesis, that it may be optimal to give no explicit incentives at all if the agent has to engage in multiple tasks and if some of these tasks cannot be contracted upon. Giving incentives for the contractible tasks can induce the agent not to engage at all in the non-contractible tasks which may be very inefficient. However, in order to show that an incomplete contract that gives no explicit incentives at all is more efficient, Holmström and Milgrom (1991) assume that the agent is intrinsically motivated and will choose sufficiently high effort levels even in the absence of any explicit incentives. In contrast, we show that, due to the existence of fair principals, it is not necessary to rely on intrinsic motivation because incomplete bonus contracts turn out to be a very effective method of efficient effort elicitation in *all* tasks. We report on an experiment that captures Holmström and Milgrom's multi-task principal agent problem. In the experiment it turned out that the incomplete bonus contract is more efficient and more profitable than a more complete piece rate contract that makes pay contingent on the effort in the contractible task.

The rest of this paper is organized as follows. Section IIA sets up a simple principal agent framework that is the baseline for the first series of experiments. In this setting there is only a single task to be performed by the agent. In Section IIB we derive alternative hypotheses that are based on the assumption that a non-negligible minority of the subjects is also motivated by fairness concerns. Section IIC describes the experimental procedures and Section IID presents the experimental results. In Part III we look at a multi-task principal agent problem in the spirit of Holmström and Milgrom (1991). Section IIIA describes the experimental design, Section IIIB derives the alternative hypotheses based on the existence of fair subjects, and Section IIIC presents the results of these experiments. Part IV summarizes our main results and concludes the paper.

## II. Contract Choice in the Presence of a Single Task

### A. The Principal-Agent Problem

This section presents a simple principal-agent problem and our first experimental design. Consider a principal who hires an agent to carry out production. Suppose that both parties are risk-neutral.<sup>3</sup> If the agent exerts effort  $e \geq \underline{e}$ , he incurs a private cost  $c(e)$  (measured in monetary terms with  $c(\underline{e})=0$ ,  $c'(e) > 0$ , and  $c''(e) > 0$ ), and he generates a gross profit  $v(e)$  that is strictly increasing and concave in  $e$ . Let  $e^{FB} > \underline{e}$  denote the unique first best efficient effort level that maximizes  $v(e)-c(e)$ .

We assume that gross profits and effort costs cannot be contracted upon. Furthermore, though both parties observe the agent's effort level, effort is only partially verifiable to the courts. For concreteness, we assume that the principal can invest in a verification technology at a fixed cost  $k$  that renders effort verifiable to the courts with probability  $p$  ( $0 < p < 1$ ). If the principal does not invest into verification, effort is not verifiable. The principal also has the option to impose a fine  $f$  on the agent. The fine has to be paid to the principal only in case of *verified* shirking because the courts enforce only those contractually specified payments that are contingent on verifiable events.  $f$  has to obey  $0 < f < f^*$  where  $f^*$  is the maximal feasible fine and  $pf^* < c(e^{FB})$ .

The timing of events is as follows. At date 0 the principal makes a take-it-or-leave-it contract offer to the agent and decides whether to incur the verification cost. The contract specifies a payment scheme for the agent. If the agent rejects the offer, both parties get their reservation utilities that we normalize to 0. If the agent accepts, he has to choose his effort level  $e$  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 the experimental design we chose the following parameters for this simple principal agent problem. The agents could choose effort  $e \in \{1, \dots, 10\}$  with effort costs given by Table 1.

---

<sup>3</sup> The assumption of risk neutrality is justified in view of the calibration results in Rabin (2000). Rabin has shown that risk aversion is highly implausible for the typical stake level prevailing in laboratory experiments. Even small degrees of risk aversion over the typical experimental stakes imply an absurd degree of risk aversion over larger stakes. Thus, if one assumes the validity of expected utility theory, one is essentially forced to assume risk neutrality over the typical experimental stake levels.

**Table 1: 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

An effort of  $e$  yields a gross profit  $v(e)=10 \cdot e$  to the principal. If the principal invests in the verification technology she can verify the agent's effort with probability  $p=1/3$  and has to incur the fixed cost  $k=10$ . The fine that can be imposed on the agent is bounded above by  $f'=13$ . 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  $v(e)-c(e) = 80$ .<sup>4</sup>

A contract which relies on effort verification and explicit, enforceable incentives, will be called *incentive contract (IC)*. An IC  $(w, e^*, f)$  stipulates a wage  $w$ , a demanded effort level  $e^*$ , and a fine  $f$ , to be paid in case of verifiable shirking. An IC implies the following (expected) monetary payoffs:

$$M^A = \begin{cases} w - c(e) & \text{if } e \geq e^* \\ w - c(e) - 1/3f & \text{if } e < e^* \end{cases} \quad \text{and} \quad M^P = \begin{cases} 10e - w - 10 & \text{if } e \geq e^* \\ 10e - w + 1/3f - 10 & \text{if } e < e^* \end{cases}$$

From the point of view of traditional contract theory, the analysis of the (second best) optimal contract is straightforward if we make the standard assumption that the principal and the agent both want to maximize their material payoffs. The agent has no incentive to provide effort above  $e=1$  except when the principal stipulates a fine and the resulting expected fine  $pf$  exceeds the agent's effort cost  $c(e)$ . For  $k=10$ , the principal will incur the verification cost and impose the maximal feasible fine  $f'=13$ . The (second best) optimal contract demands an effort level  $e^*$  from the agent such that the non-shirking condition  $pf' \geq c(e^*)$  just holds and the agent has to pay  $f'$  whenever  $e < e^*$  is verified. Given the parameters of the experiment, the optimal demanded effort level is  $e^* = 4$ . The optimal wage offer to the agent is the minimum payment ensuring that a non-shirking agent is just willing to accept the contract, i.e.,  $w = c(e^*) = 4$ . The principal's material payoff from this contract is  $M^P = ve^* - w - k = 40 - 4 - 10 = 26$  while the agent's material payoff is  $M^A = w - c(e^*) = 4 - 4 = 0$ .<sup>5</sup>

<sup>4</sup> All payoffs are in "tokens". For the conversion rate into DM see Section II.C below.

<sup>5</sup> Note that the agent is indifferent whether to accept or to reject the contract. However, in equilibrium he must accept with probability one, otherwise the principal would have been better off by offering a slightly higher wage,  $w'+\varepsilon$ , that the agent would accept for sure. However,  $w' + \varepsilon$  cannot be an equilibrium because  $w' + \varepsilon/2$  would also be accepted for sure by the agent and yields a strictly higher payoff to the principal. Thus, the only equilibrium is the one described above.



However, the above analysis rests on the important assumption that both players are interested only in their own material payoffs. To see the implications of this assumption note that effort is observable by both parties. Thus, as an alternative to the above incentive contract, the principal could simply "ask" the agent to put in  $e^* > \underline{e}$  and "promise" him a reward in return. This could be done with two different types of contracts:

- (i) *Trust Contract (TC)*: The principal offers the agent an unconditional payment  $w > c(\underline{e})$  which is enforced by the courts. In return, she asks the agent to put in effort  $e^* > \underline{e}$ , which cannot be enforced by the courts. If the agent accepts a trust contract, he is not obliged to choose  $e = e^*$ ; he can choose any feasible effort level,  $e \in \{1, \dots, 10\}$  at date 1. The monetary payoff from a trust contract  $(w, e^*)$  is given by  $M^A = w - c(e)$ , for the agent, and  $M^P = 10e - w$ , for the principal, where  $e$  is the agent's actual effort level.
- (ii) *Bonus contract (BC)*: In a BC  $(w, e^*, b^*)$  the principal offers a low, enforceable base wage  $w \geq c(\underline{e})$  and asks the agent to spend effort  $e^* > \underline{e}$ . Furthermore, the principal announces to pay a bonus  $b^* \geq c(e^*)$  if the agent chooses  $e \geq e^*$ . However, neither the agent's effort nor the principal's bonus payment are enforceable. If the agent accepts a bonus contract, he chooses any effort  $e \in \{1, \dots, 10\}$  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  $b = b^*$ ; she can choose any  $b \geq 0$ . A bonus contract implies monetary payoffs  $M^A = w - c(e) + b$ , for the agent, and  $M^P = 10e - w - b$ , for the principal.<sup>6</sup>

Both types of contract are less complete than the incentive contract because in the incentive contract the principal credibly ties compensation to the agent's performance. Moreover, in a sense the bonus contract is more incomplete than the trust contract. The trust contract is not conditional on the agent's effort, but it does bind the behavior of the principal. In a trust contract the total compensation of the agent is fixed ex-ante in a binding manner. In contrast, the bonus contract gives discretion to both parties. The agent is free how much effort to choose and the principal is free what bonus to pay.

To examine the principals' preferences for the different contracts we implemented two treatment conditions in the experiments. In the first treatment, the TI-treatment, the principals could choose between the trust and the incentive contract. In the second treatment, the BI-treatment, they

---

<sup>6</sup> In all three types of contracts the principal was constrained to choose  $w \geq c(e^*)$ . We imposed this constraint in order to make sure that loss aversion does not affect the behavior of the agents.

had a choice between a bonus and an incentive contract. Note, however, that in the BI-treatment principals could, in fact, also propose a trust contract because the trust contract is a special case of a bonus contract. When the bonus contract is available principals are free to *not* promise a bonus, i.e., instead of paying a “low” base wage,  $\underline{w}$  say, and promising a bonus,  $b^*$  say, they can always offer a binding, unconditional, wage  $w = \underline{w} + b^*$ , which is tantamount to a trust contract. Therefore, the principals’ behavior in the BI-treatment reveals their preferences with regard to all three available contract types.

Obviously, according to the self-interest model the trust contract and the bonus contract are doomed to fail. With a trust contract the agent knows that his wage is fixed and independent of his effort level. Therefore, he chooses  $e = \underline{e}$ . In a bonus contract the principal never pays the promised bonus at date 2. Anticipating this, the agent again chooses  $e = \underline{e}$  at date 1. Thus, the equilibrium wage offer in a trust or a bonus contract is given by  $w = c(1) = 0$ . This offer is accepted by the agent and payoffs are given by  $M^P = 10 - 0 = 10$  and  $M^A = 0 - 0 = 0$ . Remember that if the principal offers the incentive contract ( $w=4, e'=4, f=13$ ), her equilibrium payoff is given by  $M^P = 4 \cdot 10 - 4 - 10 = 26$  while the agent gets  $M^A = 4 - 4 = 0$ . Thus, the clear prediction of the self-interest model is that in both treatments the principal will opt for the incentive contract.

If, however, principals and agents are not only self-interested but are also motivated by reciprocity and concerns for fairness, the outcome is less clear. In the next section we will discuss the implications of fairness concerns on optimal contractual design and derive some testable predictions for our principal agent problem.

### *B. Implications of Reciprocal Fairness*

It is well known from day to day experience and from many controlled experiments that a non-negligible fraction of the people do behave in a reciprocally fair manner (see, e.g., Camerer 2000, Fehr and Schmidt 2000). In Fehr, Gächter and Kirchsteiger (1997) it has also been shown that fairness concerns are important for contract enforcement.<sup>7</sup> People who are motivated by reciprocal fairness are willing to sacrifice resources to be kind to those who are perceived to be kind (*positive*

---

<sup>7</sup> There are, however, crucial differences between our paper and Fehr, Gächter and Kirchsteiger (FGK, 1997). In FGK only *one* type of contract was feasible. FGK could thus not examine the principals’ contract preferences nor could they say anything about the efficiency of different types of contracts. Moreover, in contrast to FGK, our paper is based on a precise model of fairness. Therefore, in addition to the experimental results, we also provide precise theoretical insights into the role of fairness in incentive provision.

*reciprocity*) and to punish those who are perceived to be unkind (*negative reciprocity*). Whether an action is perceived to be kind or unkind depends – among other things - on the distributional consequences of the action.

What are the implications of the above qualitative definition of reciprocal fairness for our experiment? By offering a generous trust contract the principal can appeal to the reciprocity 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 in order to appeal to the reciprocity of the principal, and the principal may indeed reciprocate by paying a bonus voluntarily. Thus, both the TC and the BC may well be more efficient than predicted by the self-interest model. Whether they will be more efficient than the IC is, however, an open question that cannot be answered on the basis of general, qualitative, notions of reciprocity. The principal may, e.g., also appeal to the agents' reciprocity in the IC which may render the IC also more efficient than predicted by the self-interest model. Therefore, it is well possible that the IC still is the superior contract in the presence of reciprocally fair people because it allows to combine explicit material incentives with an appeal to the agents' reciprocity.

However, in the absence of a precise and tractable model of reciprocal fairness nothing specific can be said. It may be that the TC and the BC are more or less efficient than the IC. Therefore, we rely in the following on a simplified version of the theory of inequity aversion developed by Fehr and Schmidt (1999). This model captures some aspects of reciprocal fairness in a tractable way and is consistent with the outcome of many different experimental games.<sup>8</sup> However, although we use the Fehr-Schmidt model as a tool to derive quantitative predictions, we would like to stress that we do not regard our experiments as a test of this theory against other theories of fairness.<sup>9</sup>

The theory of Fehr and Schmidt (1999) has two main ingredients: First, the theory assumes that some people are not only concerned about their own material payoff but also care about inequity

---

<sup>8</sup> Fehr and Schmidt apply the theory e.g. to ultimatum, market, public good, and gift exchange games, and show that the interaction of the distribution of types and the strategic environment can explain the major facts in these games.

<sup>9</sup> In fact, it is well possible that other theories of equity (e.g. Bolton and Ockenfels, 2000) or of intention-based reciprocity (e.g., Rabin 1993, Falk and Fischbacher, 1999, Dufwenberg and Kirchsteiger, 1999, or Charness and Rabin, 2000) yield similar predictions. Our experiments have not been designed to discriminate between these models. The main reason for our application of the theory of inequity aversion is that this theory is relatively simple compared to most other fairness theories and provides nevertheless precise and interesting insights into the impact of fairness considerations on contractual choices.

or, in our context, inequality.<sup>10</sup> Second, the theory acknowledges that people differ. Some people are very much concerned about inequality and have a high willingness to pay in order to reduce it, while others only care 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 and  $\beta_i \leq \alpha_i$ ,  $0 \leq \beta_i < 1$ . In this utility function, the term weighted with  $\alpha_i$  measures the utility loss that stems from inequality to  $i$ 's disadvantage, while the term weighted with  $\beta_i$  measures the loss from advantageous inequality. We use a grossly simplified version of this theory. We assume that there are 60 percent self-interested types ( $\alpha_i = \beta_i = 0$ ) and 40 percent "fair" types. Fair subjects exhibit  $\alpha_i \geq \beta_i > 0.5$ , i.e., they have a willingness to pay in order achieve equality.<sup>11</sup> If the inequality is to their disadvantage, they are prepared to engage in costly "punishment" in order to reduce the payoff of their opponent. If the inequality is to their advantage, they are willing to spend resources in order to benefit the other player. Subjects with  $\alpha_i \geq \beta_i > 0.5$  are willing to share the surplus of a contract equally and reject offers that give them less than 25 percent of the surplus. The evidence from many experiments seems roughly compatible with the assumption that 40 percent of all subjects fall in this category.

On the basis of these assumptions, our principal agent problem can be analyzed using standard game theoretic tools. The full analysis is not difficult but somewhat lengthy and is therefore relegated to an appendix that can be found on our webpage.<sup>12</sup> Here, we want to report the main predictions that follow from this analysis and give the intuition for them:

***Hypothesis 1 [Trust contracts]:*** *Increasing the wage in a trust contract increases the effort of the fair agents, but, on average, the effort increase is too small to make a wage increase profitable for the principal.*

---

<sup>10</sup> There is no generally accepted notion of *fairness*, but probably all fairness definitions imply that equals should be treated equally. In our experiments, the subjects enter the laboratory as equals. They have no information about their opponents and do not know with whom they trade. Thus, in these very simple environments, it seems natural to define *equality* as the reference point for a fair payoff distribution.

<sup>11</sup> See Fehr and Schmidt (1999) for a more extensive discussion of the experimental evidence on the distribution of inequity averse types. When Fehr and Schmidt calibrate their model to explain the quantitative evidence in the different games they use four different types, but aggregated they also have that 40 percent of subjects exhibit  $\alpha_i \geq \beta_i > 0.5$  and that 60 percent exhibit  $0.5 > \alpha_i \geq \beta_i$ .

<sup>12</sup> Please visit: [http://www.wvl.uni-muenchen.de/l\\_s\\_schmidt/experiments/incomplete\\_contracts/index.html](http://www.wvl.uni-muenchen.de/l_s_schmidt/experiments/incomplete_contracts/index.html) .

To see the intuition for H1 consider a fair agent who accepted 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, for a fair agent  $e$  increases with  $w$ , but, if the fraction of fair agents in the population is  $q = 0.4$ , then an increase of  $w$  by 1 token increases average effort by at most  $\Delta e = 0.4 \cdot 2/11 = 0.07$  which increases the principals gross profit by at most  $10 \cdot 0.07 = 0.7$  tokens. Hence, a wage increase does not pay off for a selfish principal.<sup>13</sup> What about an inequity averse principal? A generous wage will not pay off in monetary terms, and it will generate inequality to the principal's disadvantage whenever a selfish agent chooses  $e=1$ . Hence, an inequity averse principal will not pay a higher wage either. The reason why the trust contract does not work is that the fraction of fair agents is just too small. It would have to be at least  $2/3$  in order to make it profitable for the selfish principals to offer generous wages that induce the fair agents to choose  $e > 4$ .<sup>14</sup>

**Hypothesis 2 [Incentive Contracts]:** *The optimal incentive contract stipulates the maximal fine,  $f=13$ , and demands the maximal incentive compatible effort level,  $e^*=4$ . Selfish principals offer  $w=4$ , while fair principals offer  $w=17$ , so the average wage offer is  $w=9.2$ . Selfish agents accept the contract and choose  $e=e^*$  if and only if the contract is incentive compatible. Otherwise they shirk. If the wage offer is generous ( $w \geq 17$ ), fair agents will accept and choose  $e \geq 4$ , where  $e$  is increasing with  $w$ . However, if the wage offer is greedy ( $w < 17$ ), fair agents will either reject or shirk, even if the contract is incentive compatible.*

We know already that a selfish principal would offer ( $w=4$ ,  $e^*=4$ ,  $f=13$ ) if all agents were selfish. With some fair agents, the selfish principal runs the risk that the fair agents may reject her

---

<sup>13</sup> In other environments the payment of generous wages may be profitable in a trust contract. For example, it has been shown experimentally (e.g., Fehr and Falk 1999) that if the principal's payoff function is  $M^P = (v-w)e$  with  $v = 120$  and  $e \in \{0.1, 0.2, \dots, 1\}$ , the payment of generous wages increases profits. The reason is that with this payoff function the principal's effective wage cost is  $we < w$  for all  $e < 1$ . Thus, paying generous wages is less costly compared to our experiments. The above payoff function is, however, not a natural specification for our principal agent problem.

<sup>14</sup> Note that for  $e > 4$ ,  $c'(e) \geq 2$ . 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  $q > 2/3$ , 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.

offer. Thus, she may want to increase  $w$  in order to increase the probability that her offer will be accepted. However, it turns out that with  $q=0.4$  increasing the wage does not pay off. Furthermore, increasing  $w$  somewhat, but not up to the level that aims at equality, is dangerous, because the agent may accept the contract but consider it as unfair and shirk. Therefore,  $(w=4, e^*=4, f^*=13)$  is optimal for a selfish principal. Note that with probability 0.4 the contracts of the selfish principals will be rejected. This neatly illustrates that optimal incentive contracts may become less efficient because of the existence of fair subjects. A fair principal will offer  $(w=17, e^*=4, f^*=13)$  which shares the surplus equally, if the agent chooses  $e=4$ . The maximum fine is necessary to induce the selfish agents to choose  $e=4$ . A fair principal could pay a higher wage in order to induce the fair agents to choose  $e > 4$  (which is no longer incentive compatible), but, for the same reasons as in the trust contract, this strategy will lose money on average.

Given H1 and H2 it is clear that trust contracts do not work while incentive contracts can at least be used to implement  $e=4$  in the accepted contracts. Hence, the next hypothesis is no longer surprising.

**Hypothesis 3 [TI-treatment]:** (a) *Both types of principals prefer incentive to trust contracts.*  
 (b) *Incentive contracts are more efficient and give a higher monetary payoff to the principal because they elicit, on average, a higher effort level than trust contracts.*

The analysis of the *bonus contract* is a little more complicated. The problem is, that the principal moves twice, first when he offers the contract and second when he chooses which bonus to pay. Thus, the agents may take the contract offer as a signal about the principal's type and update the probability that a bonus will be paid. However, it can be shown that no separating equilibrium exists in this signaling game and that both types of principals must offer the same bonus contract in equilibrium. This contract can be characterized as follows:

**Hypothesis 4 [Bonus contract]:** (a) *On average, principals reward higher effort levels with higher bonus payments such that non-minimal effort provision is profitable for the agents.*  
 (b) *If the a higher wage offer of the principal is not interpreted as a signal that the principal is selfish, then there exists a unique pooling equilibrium in which both types of principals offer  $w=15$ . The selfish agent chooses  $e=7$  and is rewarded by the reciprocal principal with a bonus of 25, while the selfish principal does not pay a bonus. The fair agent chooses  $e=2$  and neither type of principal pays a bonus. There does not exist a separating equilibrium.*

It is obvious that a selfish principal will not pay a bonus. A fair principal, however, pays a bonus that equalizes her monetary payoff with that of the agent:

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

Using the implicit function theorem, we get

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

If agents believe that  $q$  percent of all principals are fair and choose  $b$  in this fashion while  $(1-q)$  percent are self-interested and choose  $b=0$ , the bonus will, on average, increase with the effort level, as stipulated in H4a. 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 as compared to  $c'(e)$ . Recall that, according to the cost schedule in Table 1,  $1 \leq c'(e) \leq 4$ . For  $c'=1$ , the critical value for  $q$  is  $0.18$ , for  $c' = 2$  it is  $0.33$ , for  $c' = 3$  it is  $0.46$  and for  $c' = 4$  it is  $0.57$ . Hence, in a pooling equilibrium, where the agents believe that they face a fair principal with probability  $q = 0.4$ , profit-maximizing 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  $e = 7$ , hoping that they will be rewarded with a generous bonus payment by the principal. Fair agents choose  $e = 1$  or  $e = 2$  depending on the wage offered (for  $w \geq 10$  the fair agent chooses  $e = 2$ ). The reason for this interesting implication is that a fair agent suffers more than a self-interested agent if he meets a selfish principal who does not pay the bonus. Hence, even if it is profitable from a monetary perspective to choose  $e = 7$ , a fair agent prefers  $e = 1$  or  $e = 2$  in order to ensure that equality prevails and to avoid the disutility from disadvantageous inequality when the bonus is not paid. Thus, the presence of fair principals induces selfish agents 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.

There are many pooling equilibria in this game that differ in the unconditional base wage (and therefore also in the bonus to be paid ex post). If we impose the mild condition on out of

equilibrium beliefs that higher wage offers are not taken as a signal that the principal is more likely to be selfish, then the set of pooling equilibria shrinks to a singleton in which all principals offer  $w=15$ .<sup>15</sup> If the selfish agent chooses  $e=7$ , then a bonus of 25 just equalizes payoffs. However, this bonus is paid only by the fair principals, so the expected bonus is  $0.4 \cdot 25 = 10$ . The fair agents choose  $e=2$ , so the expected effort level is  $0.6 \cdot 7 + 0.4 \cdot 2 = 5$ .

***Hypothesis 5 [BI-treatment]: Both types of principals prefer bonus to incentive contracts. The bonus contract elicits, on average, a higher effort level than the incentive contract and both, principals and agents, achieve a higher material payoff.***

H5 is in sharp contrast to the self-interest model which predicts that the IC is much more efficient than the BC. However, in the presence of fair principals the bonus contract is more efficient and more profitable for two reasons. First, it adds additional reciprocation opportunities which are used by the fair principals to provide powerful incentives for effort provision. Note that in a pooling equilibrium the selfish principals also enjoy the high effort levels although they do not contribute to the provision of incentives. Second, because the principals can decide how much bonus to pay *after* observing the agent's effort, they can prevent to do worse than the agent. Therefore, both types of principals prefer the bonus contract to the incentive contract.

Hypotheses 1-3 look very familiar, because the equilibrium predictions are quite close to the equilibrium predictions of the self-interest model. It is, however, surprising that they are derived from a model that is based on inequity aversion. This illustrates a point of general importance:<sup>16</sup> The presence of a substantial fraction of fair types does not guarantee that the aggregate outcome diverges from the predictions of the self-interest model. However, there are also circumstances where the reverse is true: The presence of a majority of self-interested types does not guarantee that the aggregate outcome is in line with the predictions of the self-interest model. The application of the Fehr-Schmidt model to the BI-treatment illustrates this nicely. Although we still maintain the assumption that 60 percent of the subjects are completely self-interested (and rational) the predictions in H4 and H5 are very different compared the predictions of the self-interest model.

---

<sup>15</sup> Note that in a pooling equilibrium with bonus contracts wages cannot exceed 15. For  $w > 15$ ,  $b(7) = 40 - w < 25$  and, hence, the expected bonus payment is  $0.4 \cdot b(7)$  which is less than 10. Therefore, the selfish agent will no longer be willing to provide  $e = 7$ .

<sup>16</sup> In Fehr and Schmidt (1999) several other examples of this kind are provided.



### *C. Experimental Procedures*

The experiments were conducted at the University of Munich in the spring and summer of 1999. The participants were undergraduate students from the University and the Technical University of Munich (students of law, political science, engineering, etc.). The subjects were recruited with the announcement that they could earn, depending on their decisions, a considerable amount of money during the experiment.

In total we conducted four experimental sessions. Two sessions (S1 and S2) implemented the TI-treatment, the other two sessions (S3 and S4) implemented the BI-treatment. In each session we had 20-24 subjects, half of them principals ("employers") the other half agents ("employees"). The two groups were located in separate but adjacent rooms. Before the experiment started, all subjects had to read detailed instructions and to solve several exercises to make sure that all of them understand the rules of the experiment. In each session we had ten rounds. In each round an employee was matched with a different employer. Thus, we have for each subject in each experimental session ten contracts with ten different contracting partners.

After each round the subjects had to compute their own payoff and the payoff of their partner. To rule out the possibility of reputation building, the outcome of each round was strictly confidential, that is, each principal-agent pair only observed what happened in their relationship. They did not observe the contracts chosen by or offered to the other subjects in the room. Nor did they observe the past behavior of their current partner. Furthermore, the matching was random and anonymous. Finally, at the end of the session the subjects collected their total monetary payoffs privately and anonymously. Each session lasted between two and two and a half hours. A complete set of the instructions for all our experiments can be found on our webpage.<sup>17</sup>

In each session all participants received an initial endowment of DM 20.-. The experimental (token) payoffs were exchanged into money at the rate of 1 token = 0.2 DM. Thus, an employer and an employee could jointly earn a maximum surplus of DM 16 ( $\approx$ US \$ 10 at the time of the experiment) in each of the ten rounds. The highest total income of one individual was DM 115.80 (US \$72), an hourly wage of ca. DM 50 (US \$31). However, the subjects could also make substantial losses. In order to avoid the possibility that somebody ends up with negative earnings, a subject had

---

<sup>17</sup> The full set of all our experimental instructions, in German and translated into English, are available at [http://www.vwl.uni-muenchen.de/l\\_s\\_schmidt/experiments/incomplete\\_contracts/index.html](http://www.vwl.uni-muenchen.de/l_s_schmidt/experiments/incomplete_contracts/index.html).

to drop out of the experiment if his accumulated earnings fell below DM 5.- (US \$3.12), which happened three times.

#### *D. Experimental Results*

We first present the results of the TI-treatment where principals could choose between a trust contract  $(w, e^*)$  and an incentive contract  $(w, e^*, f)$ . We observed a total of 195 contractual choices in sessions S1 and S2. Nine incentive contracts and one trust contract have been rejected so that, in total, 185 effort choices have been made by the agents.

**Result 1(a):** *In the TI-treatment the clear majority of the contracts are incentive contracts and the share of incentive contracts increases substantially over time.*

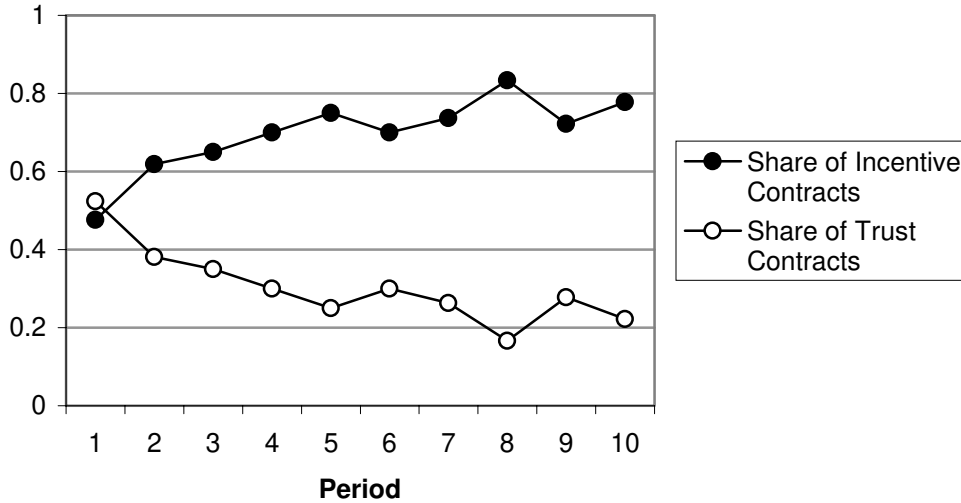
**(b)** *The average effort of the agents and the average payoff of the principals are both higher in the incentive contracts.*

Overall, 135 of the 195 contracts were incentive contracts (69 percent) and only 60 (31 percent) are trust contracts. These numbers hide, however, a strong time trend in the share of incentive contracts. Figure 1, which depicts the evolution of contractual choices in the TI-treatment, shows this trend. At the beginning slightly less than 50 percent of the proposed contracts were incentive contracts. Yet, from period 4 onwards the fraction of incentive contracts is never below 70 percent and in the final three periods roughly 80 percent of all contracts are incentive contracts. 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 they settled on the incentive contract. In view of the very high fraction of incentive contracts during the last three periods H3a receives substantial support.

R1b is illustrated by Figure 2 that depicts the evolution of average effort levels (and average desired effort levels) over time for both contracts. The figure shows that in almost all periods the average effort is higher in the incentive contracts. Moreover, in those periods in which average effort is somewhat higher in the trust contract, the fraction of trust contracts is already small, so that this is driven by very few observations. On average, the principals earned a payoff of  $-0.9$  when they proposed an incentive contract and  $-2.36$  when they proposed a trust contract. Although this result supports H3b, the support seems rather weak because, *on average*, the incentive contracts are much

less profitable and elicit lower effort levels than predicted by H2. The next result partially explains this effect.

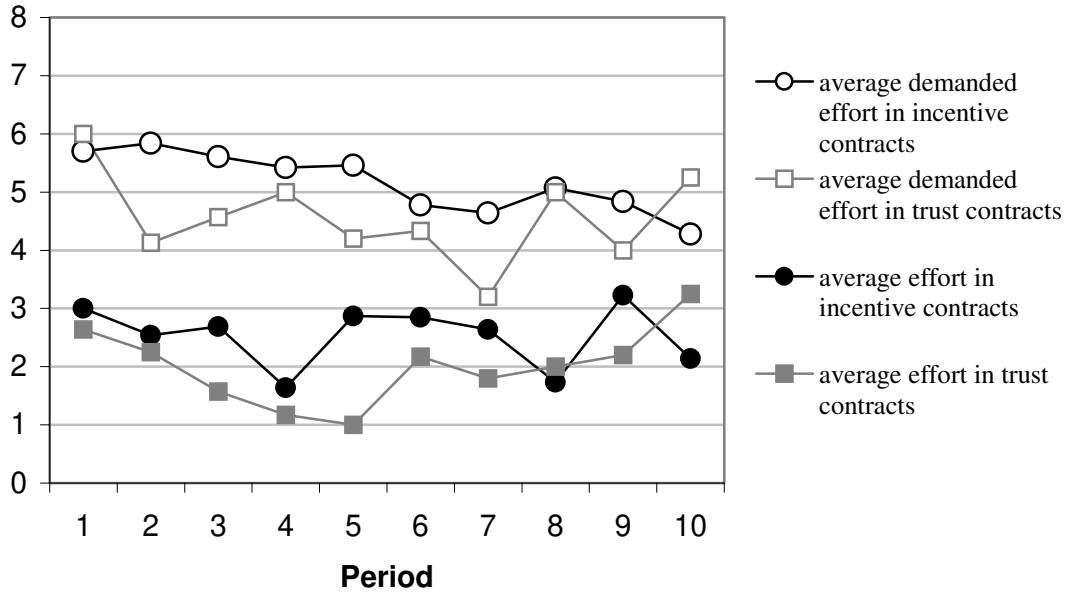
**Figure 1: Share of incentive and trust contracts (S1 and S2)**



**Result 2:** *Although most incentive contracts stipulate the maximal fine, the majority of incentive contracts violates the no-shirking condition because the principals demand too high effort levels. In addition, principals offer, on average, higher wages than predicted. There is, however, a strong increase in the share of incentive compatible contracts over time. Finally, in roughly 1/4 of all incentive compatible contracts agents shirk.*

R2 says that H2 is supported with regard to the predicted fine level but not with regard to the predicted desired effort level  $e^*$ . The average fine is 12.3. Yet, in 79 (58.5 percent) of the 135 incentive contracts the no-shirking condition is violated, i.e. principals demanded too high effort levels. This is also illustrated in Figure 2 which shows that the average demanded effort level in the incentive contracts is persistently above the maximal enforceable effort of  $e^* = 4$ . However, demanded effort levels decline over time. At the beginning the demanded effort level is, on average, around  $e^* = 6$ , in the final period it is only slightly above the predicted level of  $e^* = 4$ . This suggests that, over time, the fraction of non-incentive compatible contracts decreases, which is indeed the case. In period one, 90 percent of all incentive contracts are not incentive compatible. This fraction decreases to 36 percent in period ten.

**Figure 2: Average effort and average demanded effort in the trust and incentive contracts (S1 and S2)**



A similar time pattern is exhibited by average wages. During the first few periods, when the share of incentive compatible contracts was still low, incentive contracts stipulated on average wages well above 20. The average wage decreases, however, strongly over time and reaches a level of 11.9 in period ten which is relatively close to the predicted level of 9.2. 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 towards the predicted incentive contracts.

What was the reason for the switch to incentive compatible contracts? To answer this question Table 2 presents more detailed information on the relative profitability of the different contract types and the agents' effort behavior. The final row of Table 2 shows that the average profit in the incentive compatible contracts is 8.6 while it is  $-7.6$  in the non-incentive compatible contracts. Thus, while incentive compatible contracts are considerably more profitable than the trust contracts, the non-incentive compatible contracts are less profitable than the trust contracts.<sup>18</sup> This explains both the strongly increasing time trend in the fraction of incentive contracts (relative to the trust

<sup>18</sup> All these differences are statistically significant according to a non-parametric Mann-Whitney test ( $p < 0.005$  when the trust contract is compared to the incentive compatible contract;  $p < 0.036$  when the trust contract is compared to non-incentive compatible incentive contracts).

contract) and the strong increase in the fraction of incentive compatible contracts (relative to the non-incentive compatible contracts).

**Table 2: Wages, Effort and Principals' Payoff in the Trust-Incentive Treatment**

Wage Offer	Incentive Compatible Incentive Contracts					Non-Incentive Compatible Incentive Contracts					Trust Contracts				
	#	rej.	$e < e^*$	$e \geq e^*$	P's payoff	#	rej.	$e = 1$	$e > 1$	P's payoff	#	rej.	$e = 1$	$e > 1$	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		37	21	-2.4

Remark: In the column # the table shows the number of contract offers. The sum of the columns  $e < e^*$  ( $e = 1$ ) and  $e \geq e^*$  ( $e \geq 1$ ) gives the number of accepted contract offers.

There are two reasons why the incentive compatible contracts are more profitable than the non-incentive compatible contracts. Principals pay much higher wages when they offer non-incentive compatible contracts and, despite their generosity, shirking is much more frequent in these contracts. Table 2 shows that in all 79 offered contracts that are *not* incentive compatible the wage is above  $w = 10$  while in the majority of the incentive compatible contracts (in 29 of 56 cases) the wage is below  $w = 10$ . This suggests that the principals wanted to elicit reciprocal effort choices from the agents when they proposed non-incentive compatible contracts. Table 2 also shows, however, that these attempts frequently failed. In 48 (62 percent) of the 78 accepted non-incentive compatible contracts the agents chose the minimal effort level.<sup>19</sup> This contrasts sharply with those contracts that meet the no-shirking condition. Only in 12 (26 percent) of the 47 accepted incentive compatible contracts do the agents shirk.<sup>20</sup>

<sup>19</sup> Note that the number of accepted contracts is given by the sum of the the two effort columns. For example, for the non-incentive compatible contracts it is given by the 48 contracts with  $e = 1$  plus the 30 contracts with  $e > 1$ .

<sup>20</sup> In all 12 cases the agents shirked fully.

Table 2 also allows us to evaluate the prediction of H1, that the payment of generous wages to elicit effort levels above the incentive compatible level is not profitable for the principals. The table shows that for the trust contracts the principals' payoff is monotonically decreasing in the offered wages.<sup>21</sup> For wages below  $w = 10$  the principals earns 3.7, for wages in the middle interval ( $10 \leq w \leq 20$ ) the payoff declines to  $M^P = -1.0$  and for high wages ( $w > 20$ )  $M^P = -6.4$ . A similar relation holds for the non-incentive compatible incentive contracts. In the middle interval  $M^P = -1.4$  while for high wages  $M^P = -12.0$ . This is summarized in the following result.

**Result 3:** *As predicted by H1, increasing the generosity of the wage offer to induce non-incentive compatible effort levels decreases the principals' average payoff.*

Taken together, the theory of fairness based on inequity aversion organizes the data surprisingly well. The clear majority of contracts uses the explicit incentive option, in most cases the fine is maximal, after an initial learning period the majority of incentive contracts demands incentive compatible effort levels, and shirking is much more prevalent in trust contracts and in incentive contracts that violate the no-shirking condition. Moreover, incentive compatible contracts are much more profitable, and it does not pay for the principals to elicit reciprocal effort choices from the fair agents by paying high wages.

Since several predictions of the theory of inequity aversion are shared by the self-interest model it is interesting to ask whether there are facts in the TI-treatment that allow a joint evaluation of the two theories. There are three facts that are roughly consistent with the theory of inequity aversion but contradict the self-interest model: (i) the frequent rejection of incentive compatible contracts with wages below  $w = 10$ , (ii) the existence of many incentive compatible contracts with wages between 10 and 20, and (iii) the existence of many non-minimal effort choices in the trust contract and the non-incentive compatible incentive contracts. Table 2 shows that roughly 30 percent (8 of 29) of the incentive compatible contracts in the low-wage interval are rejected. Moreover, 46 percent (26 of 56) of the incentive compatible contracts offer wages between 10 and 20. This contradicts the self-interest model because in this model the principals appropriate the whole surplus from a contract, i. e.,  $w = c(4) = 4$ . Fair principals, in contrast, share the surplus from such a contract by offering  $w = 17$ . The fact that 46 percent of the contracts are between 10 and 20 is roughly

---

<sup>21</sup> This result is also confirmed by a simple OLS-regression of effort on wages that yields  $e = 1.08 + 0.04 w + \varepsilon$  where  $\varepsilon$  denotes the error term. The t-value for the constant is 3.11 while the t-value for the coefficient on  $w$  is 3.65. According to this regression, effort significantly increases with wages but the increase is associated with a loss: A wage increase by 10 units raises effort only by 0.4 and, hence, the expected revenue increases only by 4.

consistent with our assumption about the share of fair subjects. Finally, the self-interest model predicts minimal effort choices for the trust contract and the incentive contracts that violate the no-shirking condition. In contrast to this prediction, in 38 percent (30 of 78) of the accepted incentive contracts that violate the no-shirking condition the agents chose  $e > I$  and in 36 percent (21 of 58) of the accepted trust contracts we have  $e > I$ , too. All of these facts are consistent with H1 and H2 but contradict the self-interest model.

Next we turn to the results in the BI-treatment. As we will see the theory of inequity aversion captures the major facts in this treatment even better than in the TI-treatment while the self-interest model does worse than in the TI-treatment and is clearly rejected by the data. In total we observed 230 contract offers in the BI-treatment. Four bonus contracts and two incentive contracts were rejected so that we have 224 accepted contracts in the BI-treatment. Our first result is again about the principals' contract preferences.

**Result 4:** (a) *In the BI-treatment the overwhelming majority of all contracts are bonus contracts.*  
(b) *The average effort and the average payoff of the principals is much higher in the bonus contract as compared to the incentive contract.*

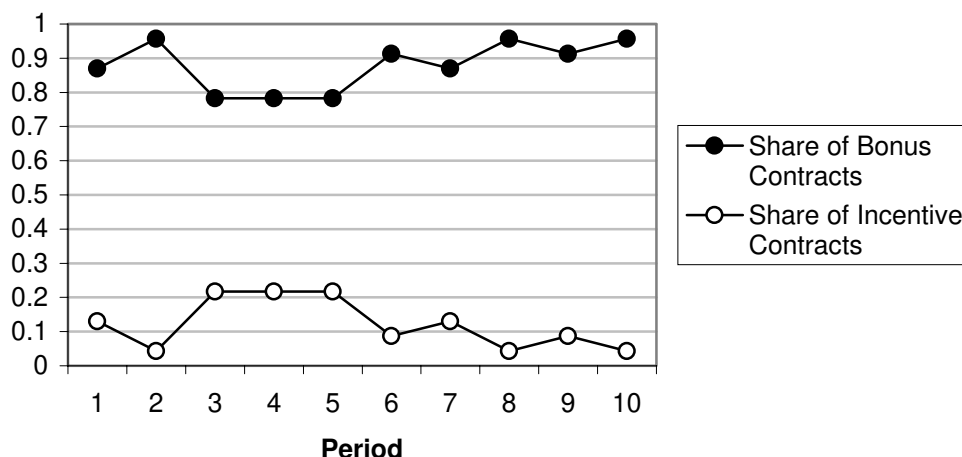
Result 4a and b provide unambiguous support for H5 and constitute an equally unambiguous rejection of the self-interest model. Figure 3 presents the evidence in favor of R4a. It shows the evolution of the share of bonus contracts over time. Already in period one 87 percent of all contracts are bonus contracts. 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.<sup>22</sup> Yet, from period six onwards the share of bonus contracts is roughly 90 percent and approaches even 96 percent in the final period. There can thus be little doubt that principals strongly prefer the bonus contract.<sup>23</sup>

---

<sup>22</sup> 57% of all principals tried an incentive contract at least once, but only 4% did so more than three times.

<sup>23</sup> In the BI-treatment trust contracts, i.e. contracts in which the principals promise and pay no bonus, are completely absent.

**Figure 3: Share of bonus and incentive contracts (S3 and S4)**

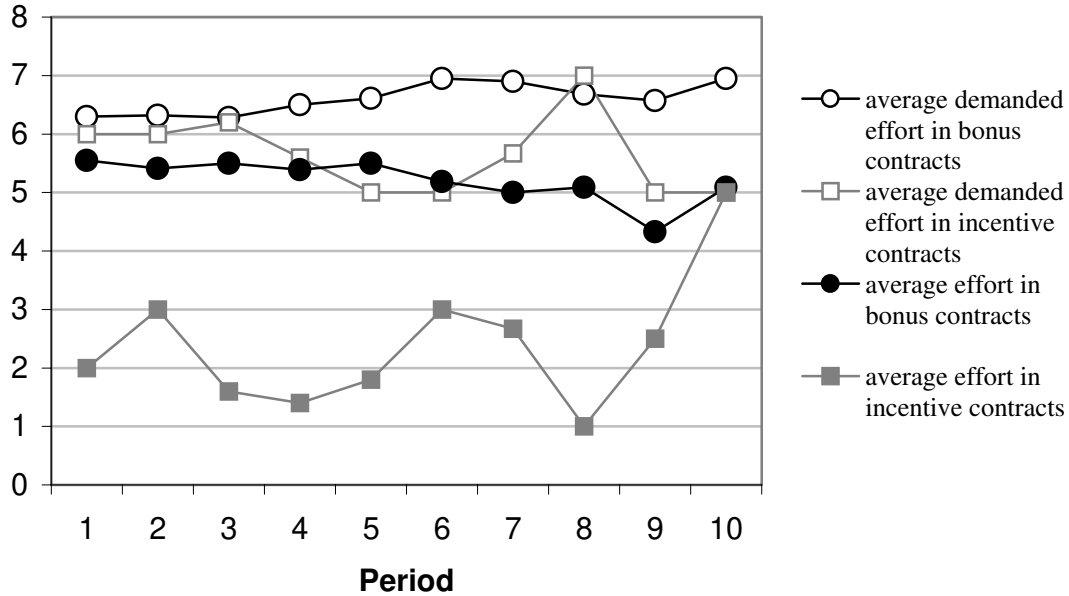


To examine the reasons for this preference we compare the average effort level in bonus and incentive contracts (see Figure 4). The figure shows that the average effort is considerably higher in the bonus contracts in nine out of ten periods.<sup>24</sup> While in the incentive contracts the average effort is, in general, between  $e = 2$  and  $e = 3$ , effort in the bonus contracts is, in general, above  $e = 5$ . Figure 4 also indicates that agents' effort in the bonus contracts is somewhat below the desired effort level but the gap between actual and desired 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 desired average effort is always above  $e^* = 4$ . 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 27 tokens while the incentive contract generates an average *loss* of 9 tokens. In each of the ten periods the average profit from bonus contracts is *always* above 20 tokens while in six of the ten periods the incentive contract causes losses and in the remaining periods the profits are, in general, rather low. In view of these large profit differences it is no longer surprising that principals exhibit a strong preference for bonus contracts.

<sup>24</sup> The exemption is period ten where the effort difference is negligible. Furthermore, in the incentive contracts the average effort in period ten is based on a single data point.



**Figure 4:** Average effort and average demanded effort in the bonus-incentive treatment (S3 and S4)



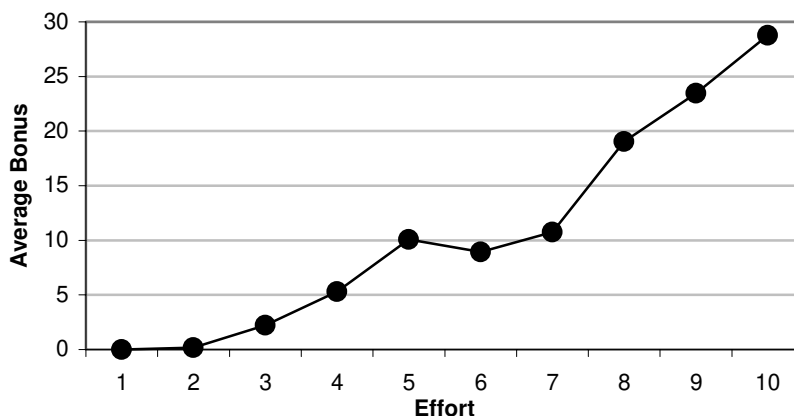
The higher effort level in the bonus contracts imply a higher surplus. To what extent did the agents receive part of this increase in the surplus relative to the incentive contracts. On average agents earned an income of 14.4 in the incentive contracts while in the bonus contracts their payoff was 17.8. This means that agents received a small part of the surplus increase while the bulk of the increase is reaped by the principals. This shows that the option to pay a bonus yields a substantial efficiency increase *and* causes sizable changes in the distribution of the surplus.

Why is the bonus contract doing so much better than the incentive contract? According to H4 the predicted reason for the superiority of the bonus contract is that principals respond reciprocally to the agents' effort choice. Our next result shows that this is indeed the case:

**Result 5: (a)** *The average bonus is strongly increasing in the effort level so that non-minimal effort choices are profitable for the agents.*

**(b)** *Principals offer on average wages of 15 and pay an average bonus of 10.4 while the agents average effort is equal to 5.2.*

**Figure 5:** Bonus-effort relation in the bonus-incentive treatment (S3 and S4)



Evidence for R5a is presented in Figure 5 which shows the average bonus as a function of the actual effort level. The figure exhibits a strongly increasing bonus-effort schedule. If the agents provide minimal or close to minimal effort levels the bonus is indeed zero while for high effort levels the average bonus approaches  $b = 30$  tokens. The positive slope of the bonus-effort schedule is also confirmed by the following regressions:

$$b = \alpha_0 + \alpha_1 e + \alpha_2 e^* + \alpha_3 w + \alpha_4 b^* + \varepsilon$$

where  $\varepsilon$  represents the error term. Table 3 reports the outcome of this regression for each session and for S3 and S4 together.

**Table 3:** Determinants of bonus payments

Session	$N$	$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$Adj. R^2$
S3	85	-10.36 (-3.71)	3.06 (10.39)	1.16 (1.80)	-0.43 (-3.22)	0.15 (1.65)	0.68
S4	113	-1.51 (-0.57)	2.69 (9.84)	-0.01 (-0.01)	-0.24 (-1.59)	0.04 (0.54)	0.50
S3-S4	198	-5.58 (-2.97)	2.86 (14.23)	0.33 (0.87)	-0.30 (-2.98)	0.12 (2.15)	0.57

Note:  $N$  denotes the number of observations, t-values are in parentheses.

All regressions indicate a sizable and significantly positive effort ( $\alpha_1$ ) coefficient. On average, an increase in the effort level by one unit increases the expected bonus payment by 2.86 tokens. Note that this 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 relation – exactly as predicted by H4. The impact of the desired effort level ( $\alpha_2$ ) is small and not significant, suggesting 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. This is roughly in line with the predictions of the theory of inequity aversion because only fair principals reduce the bonus by one unit for every wage increase by one unit (as long as  $w \leq 15$ ). Thus, with 40 percent fair principals the predicted wage coefficient in the bonus regression is 0.4. The announced bonus enters significantly with a positive, but very small coefficient. 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 very much committed to their bonus announcements and that the effort level is the major determinant of the principals' bonus choice.

Even though principals do reciprocate quite strongly on average, it is important to notice that there are big differences in individual behavior. Considering only those (162) contracts where the agents chose a non-minimal effort level ( $e > 1$ ), the principals did not pay any bonus in 34 cases (21%). Among those who did pay a bonus, many paid very little even if a high effort level was chosen. However, there were also some principals who reciprocated very generously to high effort levels. These observations are only partially consistent with the theory of inequity aversion. The problem is that our model assumes (for simplicity) that there are only two types of individuals, those who do not pay any bonus and those who pay a very high bonus that equalizes payoffs. The experimental evidences shows that many subjects were in between these extreme cases.

To what extent does the theory of inequity aversion predict the average values of  $w$ ,  $b$ , and  $e$  correctly? Recall from H4b that the theory predicts a wage of 15, an average bonus of 10 and an average effort of 5. This is almost exactly the same as reported in R5b. Thus, despite our simplifying assumptions, the theory predicts the average behavior in the BI-treatment surprisingly well.

Taken together the TI- and the BI-treatment show that the principals strongly prefer the bonus contract. If this contract is not available they prefer the incentive over the trust contract. In terms of the average effort and the average surplus associated with the three types of contracts the same ranking holds. Note, that this ranking implies a non-monotonic relationship between performance and contractual incompleteness. The bonus contract is the most incomplete contract. The agent is free which effort to choose and the principal is free what bonus to pay. The trust contract is

somewhat less incomplete. Here the agent is still unconstrained in his effort choice, but the total compensation is fixed in advance and is not conditional on the effort level. The most complete contract is the incentive contract which restricts the principal and the agent. Our experiment shows, that the most incomplete contract induces the highest effort and generates the highest profit for the principals, while the less incomplete trust contract does slightly worse than the complete incentive contract.<sup>25</sup>

### **III. Contract Choice in the Presence of Multiple Tasks**

In the previous section the agents had to perform a single task only. While this is a convenient simplification it is clear that in reality most agents have to perform multiple tasks. When there are multiple tasks it is often the case that some aspects of the performance of an agent can be contracted upon, while other aspects can not. In the example of the university professor, it may be relatively easy to measure (and contract upon) the “quantity of production”, i.e. the number of pages written in academic journals, the number of hours taught, the number of PhD students supervised, etc., but it may be much harder to measure the “quality” of production. Holmström and Milgrom (1991) argue that in such an environment it may be optimal to give no incentives at all. The reason is that if the principal rewards the measurable performance in one task, then the agent will shift too much effort into this measurable activity and disregard the other tasks that are not rewarded even though they are very important. However, if no incentives are offered, then, under the standard assumptions that agents are selfish and dislike effort, the agent will put in the minimum level of effort in both tasks.

In order to show that a trust contract can outperform an incentive contract, Holmström and Milgrom (1991) assume that the agent’s job is sufficiently *intrinsically* interesting and rewarding. Therefore, the agent will spend some positive amount of effort even in the absence of any incentives, and, because there is no distorting piece rate, he will allocate his effort efficiently across tasks. In the absence of intrinsic motivation the trust contract is, however, not viable. The previous experiments also illustrate the relatively bad performance of the trust contracts. Remember that the principals’ payoff was very low and frequently even negative in the trust contracts. This non-viability of trust

---

<sup>25</sup> Recall that the principal had to pay a fixed verification cost of 10 if he chose the incentive contract. Without this cost, the incentive contract would do considerably better than the trust contract, but still much worse than the bonus contract.

contracts in the absence of sufficient intrinsic motivation seems to severely limit the multi-tasking argument for incomplete contracts.

The situation is, however, very different if, instead of trust contracts, bonus contracts are considered. Because of the presence of fair principals bonus contracts, although highly incomplete, may provide powerful implicit incentives for effort provision. Moreover, bonus contracts may avoid an inefficient effort allocation across tasks because the actual bonus can be made dependent on the performance of the agent in *all* tasks. Therefore, if incomplete bonus contracts are taken into account, the multi-tasking argument for incomplete contracts may be valid irrespective of whether agents are intrinsically motivated, i.e., in a much larger class of environments. To test this conjecture we have conducted multi-task principal agent experiments in which effort costs are strictly increasing for *all* effort levels. Therefore, self-interested agents will never provide more than the minimal effort if they face no material incentives.

### *A. The Multi-Tasking Experiment*

In this experiment the principal can choose between a piece rate contract and a bonus contract. We call this treatment, therefore, piece rate-bonus (PB) treatment. Note that the trust contract can be considered as a special case of a “bonus” contract in which the principal chooses  $b^* = 0$ . In contrast to the previous treatments there are two separate tasks and the effort levels in these task are denoted by  $e_1$  and  $e_2$ . Effort is observable in both tasks but only  $e_1$  is verifiable and can be contracted upon. For example,  $e_1$  could measure the quantity of production and  $e_2$  its quality. Or,  $e_1$  may be interpreted as effort invested in production and  $e_2$  as effort invested in maintaining the machinery. In both of these examples it may be easy to verify  $e_1$  but prohibitively costly if not impossible to verify  $e_2$ . In both tasks the set of feasible effort choices is given by  $e_i \in \{1, \dots, 10\}$ , and the total effort cost is  $c(e_1+e_2)$ , where  $c(\cdot)$  is strictly increasing according to the cost function presented in Table 4. Note that  $e_1$  and  $e_2$  are perfect substitutes in the agent's cost function. Gross profit of the principal is given by  $v(e_1, e_2) = 10 \cdot e_1 \cdot e_2$ . Thus, efforts are complements at the margin from the point of view of the principal.

**Table 4:** Effort cost function in the piece rate-bonus treatment

$e=e_1+e_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$c(e)$	10	15	20	25	30	35	40	45	50	60	70	80	90	100	110	120	130	140	150

At date 0 the principal can offer a piece rate contract, denoted by  $(e_1^*, e_2^*, w, s)$ , that specifies a desired effort level for each task,  $(e_1^*, e_2^*)$ , a binding wage,  $w$ , and a binding piece-rate  $s$  (sharing rule) for every unit of actual effort  $e_1$ . Alternatively, the principal can offer a bonus contract, denoted by  $(e_1^*, e_2^*, w, b^*)$ . This contract also specifies a desired effort level for each task,  $(e_1^*, e_2^*)$ , and a binding base wage,  $w$ , but instead of the piece-rate it announces a non-binding bonus,  $b^*$ .<sup>26</sup> Both parties know that the bonus payment cannot be enforced and is completely voluntary. At date 1 the agent accepts or rejects the contract offer. If he rejects, both parties get a payoff of zero. If he accepts, he chooses his actual effort levels,  $(e_1, e_2)$ . At date 2, contracts are executed and, if a bonus contract has been chosen, the principal decides how much bonus to pay. Payoffs are given by

$$M^A = \begin{cases} w + se_1 - c(e_1 + e_2) & \text{in case of a piece rate contract} \\ w + b - c(e_1 + e_2) & \text{in case of a bonus contract} \end{cases}$$

$$M^P = \begin{cases} 10e_1e_2 - w - se_1 & \text{in case of a piece rate contract} \\ 10e_1e_2 - w - b & \text{in case of a bonus contract} \end{cases}$$

The predictions of the self-interest model for the PB-treatment are straightforward. In any piece rate contract a self-interested agent chooses his effort in the non-verifiable task to be  $e_2 = 1$ . Hence, the marginal revenue of  $e_1$  is 10. Given  $e_2 = 1$ , the agent's marginal cost of effort in task 1 is 5 if  $e_1 \leq 9$  and 10 if  $e_1 > 9$ . Therefore, the profit-maximizing piece rate contract is to offer  $w = 5$ ,  $e_1^* = 9$ , and  $s = 5$ , which induces the agent to accept the contract and to choose  $e_1 = 9$  and  $e_2 = 1$ . Hence, the principal's monetary payoff from a piece rate contract is  $M^P = 10 \cdot 9 \cdot 1 - 5 - 5 \cdot 9 = 40$  while the agent gets  $M^A = 5 - c(9+1) + 5 \cdot 9 = 0$ .

With a bonus contract a rational agent foresees that the bonus will never be paid and chooses  $e_1 = e_2 = 1$ . Thus, the principal should offer  $w = c(1+1) = 10$ , while the announced bonus and the desired effort levels are cheap talk. In equilibrium the agent accepts this offer, the principal's

<sup>26</sup> In both contracts the binding part of the compensation had to cover the cost of the desired effort levels. That is, the bonus contract requires  $w \geq c(e_1^* + e_2^*)$ , while the piece rate contract requires  $w + se_1^* \geq c(e_1^* + e_2^*)$ . As in the TI- and BI-treatments we imposed these constraints to guarantee that loss aversion of the agents does not affect the results.

monetary payoff is  $M^P = 10 \cdot 1 \cdot 1 - 10 = 0$  and the agent also gets  $M^A = 10 - c(1+1) = 0$ . The maximum surplus in the PB-treatment is achieved if the agent chooses  $e_1 = e_2 = 10$ , which yields  $M^P + M^A = 10 \cdot 10 \cdot 10 - 150 = 850$ .

These predictions of the self-interest model imply that the principals prefer the piece rate contract – despite the presence of multiple tasks and despite the fact that the principals can only provide incentives for  $e_1$ . The reason is that (in contrast to Holmström and Milgrom) the principal cannot increase performance in task 2 by reducing the incentives for task 1. A self-interested agent will always choose the minimal effort in task 2, irrespective of the incentives provided for task 1. Hence, according to the self-interest model there should be no interaction between the two tasks in our experiment.

### *B. Hypotheses*

What are the predictions if there is a fraction of inequity averse subjects? We rely here again on the model by Fehr and Schmidt and maintain all the previously made assumptions regarding the characteristics and the share of fair types.

***Hypothesis 6:*** *Principals in the PB-treatment prefer the bonus contract relative to the piece rate contract.*

***Hypothesis 7:*** *(a) Piece rate contracts are associated with a very uneven, i. e., inefficient, allocation of effort across tasks while in the bonus contracts the effort allocation is efficient. (b) Moreover, total effort and the material payoff of the principals is higher in the bonus contracts.*

***Hypothesis 8:*** *On average the bonus payments are increasing in total effort and decreasing in effort differences across tasks. Thus there is an implicit incentive for the agent's to provide more than the minimal effort and to equalize effort levels across tasks.*

***Hypothesis 9:*** *If a higher wage offer of the principal is not interpreted as a signal that the principal is selfish, then there exists a unique pooling equilibrium in which both types of principal offer  $w = 225$  and principals pay, on average, a bonus of  $b = 84$ . On average, agents choose an effort level of  $e_1 = e_2 = 8.4$ . There does not exist a separating equilibrium*

Again, it is not very difficult to solve this game with Fehr and Schmidt's model of inequity aversion, but it is a little bit involved and the complete analysis can be found on our webpage (see Footnote 16) . However, the intuition for these hypotheses is very similar to the intuition we provided for hypotheses H4 and H5. It does not pay for the principal to appeal to the agent's fairness in the piece rate contract because there are too few fair agents. Therefore, it is not possible to induce high effort levels in task 2 with a piece rate contract; most agents will engage only in task 1. In contrast, with a bonus contract, the share of fair principals is sufficient to provide strong pecuniary incentives for the selfish agents in *both* tasks. Since fair principals will pay the bonus at date 2 only if performance is to their satisfaction, they can penalize low total effort *and* an uneven allocation of effort across tasks. As a consequence the agents perform well in both tasks which renders the bonus contract more profitable.

### *C. Results*

The PB-treatment consists of three experimental sessions. In each session we had again 20-24 subjects, half of them principals, the other half agents.<sup>27</sup> In total there are 330 contract offers in the PB-treatment. 7 piece rate contracts and 6 bonus contracts were rejected which makes a total of 317 accepted contract offers and effort choices.

***Result 6:*** *In the PB-treatment a large majority of all contracts are bonus contracts.*

Support for R6 is provided in Figure 6 that shows the time pattern of the share of bonus contracts. In period 1 the share of bonus contracts is already above 70 percent and it never falls below that level. In some periods the principals experimented somewhat with the piece rate contract but only 27 percent of the principals offered a piece rate contract more than three times. After period 1 the share of bonus contracts fluctuates between 75 and 88 percent and in the final period it reaches 85 percent. If one averages over all periods the share is 81 percent. Thus, the bonus contract clearly is the

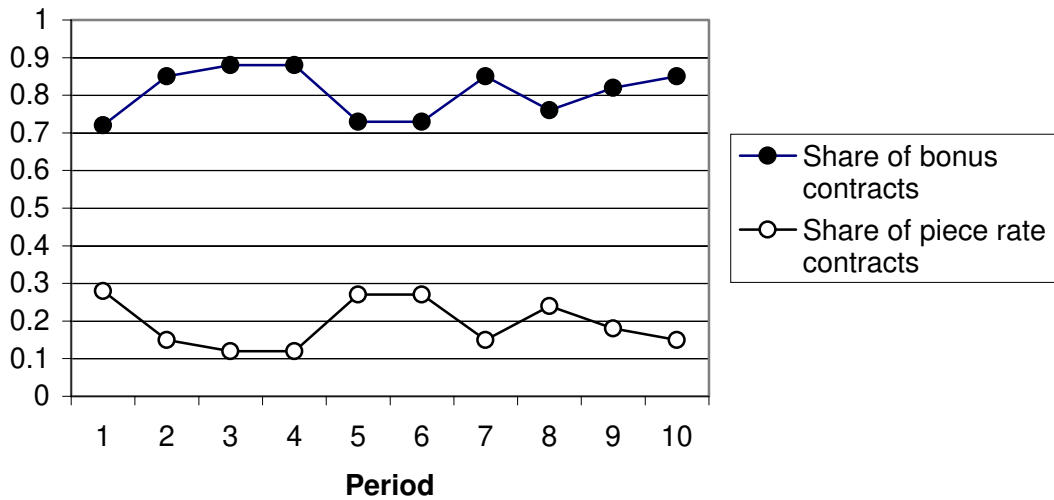
---

<sup>27</sup> In the PB treatment we applied the same experimental procedures as in the TI- and the BI-treatment (see Section II.B). The only difference was that 1 token = DM 0.02 instead of DM 0.2. This change was required because the revenue function in the PB treatment is given by  $10e_1e_2$  whereas in the TI- and the BI-treatment it was  $10e$ . The maximal revenue was thus exactly ten times higher in the PB-treatment. Subjects' average earnings in the PB-treatment in DM are similar to those in the BI-treatment. A full set of instructions in German and translated into English is available at [http://www.vwl.uni-muenchen.de/ls\\_schmidt/experiments/incomplete\\_contracts/index.html](http://www.vwl.uni-muenchen.de/ls_schmidt/experiments/incomplete_contracts/index.html) .



preferred choice by the principals which confirms H6.<sup>28</sup> The reason for the relative popularity of the bonus contract can be seen from our next result.

**Figure 6:** Share of bonus and piece rate contracts (S5 - S7)



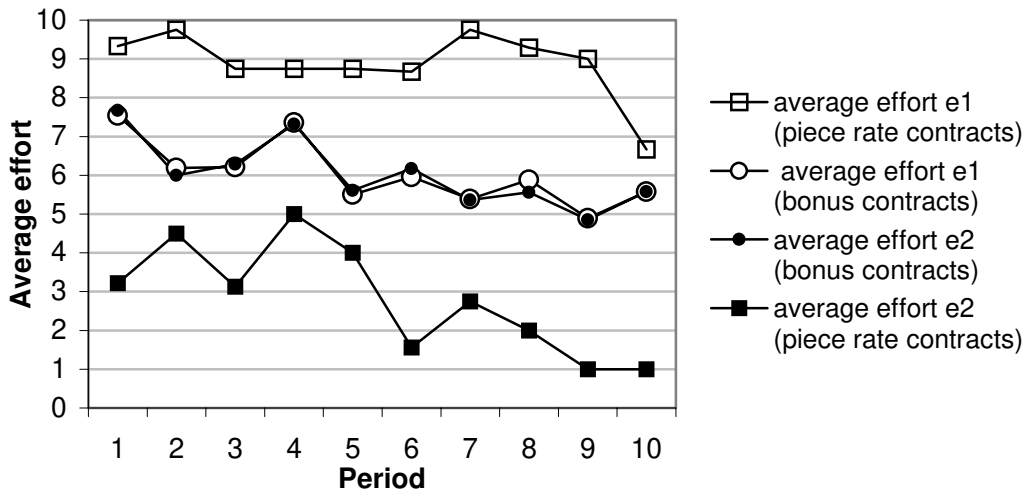
**Result 7:** (a) In the piece rate contracts the effort allocation is very uneven across tasks while in the bonus contracts the effort levels are, on average, nearly identical in both tasks. (b) The total average effort in the bonus contract is roughly equal to the total average effort in the piece rate contract. (c) Due to the much more efficient effort allocation across tasks the principals’ payoff is much higher in the bonus contracts.

Figure 7 shows the evidence for R7a. It depicts the evolution of the average effort levels in both tasks and for both contracts. The figure exhibits three features: First, in the bonus contracts the average effort across tasks is nearly indistinguishable. There is also a slight decrease in the effort over time in these contracts. Second, in the piece rate contracts the effort in the verifiable task,  $e_1$ , is higher than the effort levels in the bonus contracts. Only in period 10,  $e_1$  is substantially below 9. Finally, the effort in the non-verifiable task,  $e_2$ , is much lower than in the bonus contracts. In the final two periods  $e_2$  converges, as predicted, to the minimal level. Thus, Figure 7 provides strong support for H7a, the bonus contract achieves an efficient effort allocation while the piece rate contract does

<sup>28</sup> Trust contracts, i.e. contracts in which the principals promise and pay no bonus, are completely absent.

not. However, H7b, which stipulates a higher total effort in the bonus contracts, is only met in the final period. In the other periods total effort  $e_1 + e_2$  is similar in both types of contract.

**Figure 7: Average effort in piece rate and bonus contracts (S5 - S7)**



Because average graphs frequently hide differences at a more disaggregated level it is also useful to take a closer look at the effort distributions in the bonus and the piece rate contracts (see Table 5). The table shows that in the bonus contracts most effort choices are close or at the main diagonal which indicates efficient effort allocations across tasks. In addition, the effort distribution is bimodal. 25 percent of all observations (66 of 261) are at the minimal effort levels (1,1) and 34 percent (89 of 261) are at effort levels bigger or equal to 9 for both tasks. In contrast, in the piece rate contracts it is extremely rare that the agents chose the minimal effort in the verifiable task which illustrates the effectiveness of the piece rate in the elicitation of  $e_1$ . However, Table 5 also shows that the most frequent effort choice in the non-verifiable task is  $e_2 = 1$ . In fact, in 26 of 57 (46 percent) of all piece rate contracts this is the case. Thus, Table 5 also indicates the highly inefficient effort allocation in the piece rate contracts.

**Table 5:** Effort distribution in bonus and piece-rate contracts

Effort distribution in bonus contracts											Effort distribution in piece-rate contracts											
$e_1/e_2$	1	2	3	4	5	6	7	8	9	10	$e_1/e_2$	1	2	3	4	5	6	7	8	9	10	
1	66	1									1	1										
2				1																		
3	1	1	3	1	1		1															
4		1	1	6	2			1		1												
5				1	16	5											1					
6					4	7	4															
7						1	12	6														
8					1	1	6	10	2	2												
9	1						2		20	2												
10					1			2	4	63												

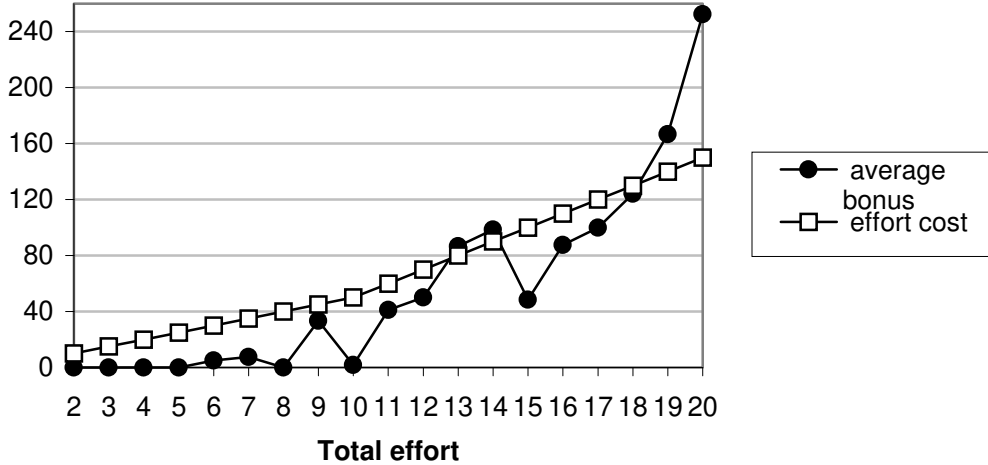
Due to the complementarity of the effort in the revenue function, principals' revenues and, hence, their profits are much higher in the bonus contracts. In all ten periods the principals' average profit in the bonus contracts is higher than in the piece rate contracts. If we average over all ten periods, the principals earned 237 tokens in a bonus contract while their income from a piece rate contract is only 41 tokens. Moreover, in four out of ten periods the principals average profit is even negative in the piece rate contracts while in the bonus contracts it is never below 178 tokens. A similar picture arises with regard to the agents' material payoff. In each of the ten periods the agents earn substantially higher incomes in the bonus contracts than in the piece rate contracts.

In view of the unambiguous payoff advantage of the bonus contract it becomes clear why the principals preferred this contract. It is, so far, however, less clear why the bonus contract was so successful in establishing an efficient effort allocation. The answer to this question is provided by our next result.

**Result 8:** *The average bonus is increasing in total effort and decreasing in effort differences. It was profitable for the agents to provide non-minimal total effort and equalize the effort across tasks in the bonus contracts.*

The first piece of evidence for R8 is given in Figure 8 which shows the actual average bonus as a function of total, actual effort.

**Figure 8:** Average bonus and effort cost as a function of total effort (S5 - S7)



The figure shows that total effort levels below 10 were almost never rewarded with a bonus. For total effort levels above 10 the bonus-effort relation is, however, increasing and the principals made substantial bonus payments. For the maximum effort level, e. g., the agents received a reward of roughly 250 tokens. To examine whether it was profitable for the agents to choose non-minimal effort levels Figure 8 also depicts the effort cost function. The comparison between the effort cost function and the bonus-effort relation reveals that it was, on average, not profitable for the agents to choose non-minimal total effort levels below 18. For the two highest effort levels, the agents received, however, sufficiently high rewards to render the choice of these effort levels profitable. The choice of the maximal effort level, in particular, turns out to be very profitable for the agents. Figure 8 also provides an explanation for the fact that the majority of the effort choices were either minimal or maximal. This was perfectly rational given the pattern of bonus payments.

To further examine the determinants of bonus payments we ran the OLS-regression

$$b = \alpha_0 \cdot d + \alpha_1 \cdot (1 - d) + \alpha_2 \cdot (e_1 + e_2) \cdot d + \alpha_3 \cdot (e_1 + e_2) \cdot (1 - d) + \alpha_4 \cdot |e_1 - e_2| + \alpha_5 \cdot w + \alpha_6 \cdot b^* + \varepsilon,$$

where  $d$  is a dummy variable that takes on the value of one if total effort is below 10. Thus  $\alpha_2$  measures the slope of the bonus-effort relation for  $e_1 + e_2 < 10$  and  $\alpha_3$  measures the slope for  $e_1 + e_2 \geq 10$ . We also included the absolute value of  $e_1 - e_2$  as a regressor to examine whether principals penalized unequal effort levels across tasks. In addition, the regression measures the impact of the

base wage and the promised bonus on actual bonus payments. Table 6 presents the results of our regressions.

**Table 6:** Determinants of bonus payments in the piece rate – bonus treatment

Variable	Session 5	Session 6	Session 7	S5 – S7
$\alpha_0$	-44.34 (-41.18)	11.23 (16.21)	4.82 (51.37)	0.19 (20.40)
$\alpha_1$	-170.23 (69.67)	-49.64 (39.25)	-129.05 (64.58)	-225.45 (30.90)
$\alpha_2$	-0.95 (9.65)	0.78 (2.88)	17.56 (9.71)	3.16 (3.92)
$\alpha_3$	20.61 (3.64)	6.36 (2.60)	19.37 (3.32)	23.09 (1.87)
$\alpha_4$	-13.47 (8.73)	-0.56 (5.15)	-60.81 (26.39)	-13.09 (5.76)
$\alpha_5$	-0.01 (0.18)	-0.06 (0.08)	0.12 (0.14)	-0.05 (0.07)
$\alpha_6$	0.18 (0.09)	-0.02 (0.04)	-0.25 (0.13)	0.01 (0.04)
$R^2$	0.62	0.13	0.76	0.74
$N$	107	79	75	261

Note: Standard errors are in parentheses,  $N$  denotes the number of observations

Most importantly, the regressions show that, for  $e_1 + e_2 \geq 10$ , the bonus rises with total effort in all sessions. The  $\alpha_3$ -coefficient is always relatively high and estimated with a small standard error. The second important results is that the  $\alpha_4$ -coefficient is negative in each session and for the data of all sessions, i.e., uneven effort allocations across tasks are penalized by the principals.<sup>29</sup> Thus, agents had a good reason to equalize effort levels across tasks. Since  $\alpha_5$  is not significant and very small the base wage does not seem to affect the bonus. The promised bonus is also not significant except in session 5 which means that, on average, the promises are just cheap talk. However, with regard to the impact of  $b^*$  on  $b$  the average is misleading because there are significant differences across individuals. If we add to the regressions in Table 6 interaction terms between  $b^*$  and the dummies for

the individual principals, significant differences among the principals can be observed. For 26 percent of the principals  $b^*$  has a significant positive impact on  $b$  – usually in the range between 0.3 and 0.5. Thus, these principals seem to feel some obligation to do what they promised. Finally, we turn to the numerical predictions of our model as summarized in H9.

**Result 9:** *The average wage and the average effort in both tasks is lower than predicted while the average bonus is higher than predicted in Hypothesis H9.*

Remember that we predicted an average wage of 225. In fact, the average wage fluctuates across periods between 120 (the minimum) and 172 (the maximum). In this context it is worthwhile to mention that our wage prediction is based on the plausible “refinement” that, in equilibrium, the agents’ belief of facing a fair principal is nondecreasing in the offered wage. In the absence of this assumption our model does not give us a unique wage prediction. However, while this refinement worked very well in the BI-treatment, it does not deliver the right predictions in the PB-treatment. With regard to the average effort the model is also not completely accurate because in most periods effort in both tasks is roughly 6 instead of the predicted value of 8.4. In comparison to the wage prediction, our bonus prediction is much better. We predicted an average bonus of 84 while the actual average bonus is 95.

Although our model does not yield accurate predictions in all dimensions it captures the major regularities of the PB-treatment quite well whereas the self-interest model is clearly rejected by the data: The bonus contract is more efficient and yields higher incomes for the principals and the agents. Principals predominantly choose the bonus contract and a substantial fraction of the principals rewards high effort and penalizes inefficient effort choices. It is thus not necessary to rely on intrinsic motivation to explain the use of incomplete contracts in the context of multiple tasks. The incomplete bonus contract is superior to the more complete piece rate contracts even in the absence of any intrinsic motivation.

---

<sup>29</sup> The  $\alpha_4$ -coefficient is not significant in S5 and S6. The likely reason for this is that in these sessions there are only a few observations with unequal effort allocations across tasks. This interpretation is supported by the fact that in S5-S7  $\alpha_4$  is highly significant.

## IV. Concluding Remarks

In this paper we have shown theoretically that the existence of fair people may have important consequences for the optimal provision of incentives. Incentive contracts that are optimal when there are only selfish actors perform less well when some agents are concerned about fairness. On the other hand, incomplete bonus contracts that cannot work when all actors are selfish provide powerful incentives and become superior when there are also fair-minded players. These predictions receive considerable support by our experiments. In particular, the principals seem to understand that fairness matters and predominantly choose the contract predicted by the fairness approach. Our theoretical and experimental results also indicate that the presence of fair players may be an important reason for contractual incompleteness. The incomplete bonus contract uses reciprocity as an enforcement device. It gives more discretion to the principals to reciprocate to the agent's effort, which renders this contract more efficient than more complete contracts that rely on explicit incentives only.

There are several other points that deserve to be emphasized. First, it is important to keep in mind that only some subjects are "reciprocators" who are prepared to reward friendly behavior and to punish unfriendly behavior even if this is costly to them. There is also a considerable fraction of subjects who seem to be interested only in their own material payoff. Second, whether or not appealing to reciprocal fairness works as an enforcement device depends on the fraction of fair types in the population and on the strategic situation in which the subjects interact. We have seen, for example, that the trust contract is not a good alternative to the more complete incentive contract. Third, we have shown that the multi-task argument for contractual incompleteness holds under more general conditions than envisaged by Holmström and Milgrom (1991). Although under the standard assumption of purely self-interested behavior the piece rate contract is predicted to be the preferred choice of the principals, the clear majority of the principals prefers the incomplete bonus contract. Fourth, our theoretical results show that simple and tractable models of fairness can yield interesting and non-obvious insights into the problems of incentive provision. The Fehr-Schmidt (1999) model turns out to be surprisingly accurate in explaining the details of our experimental results and it organizes the qualitative data pattern in all treatments very well.

Our theoretical analysis and our experiments also 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 "exploitative"

situations in which the other party may take advantage of them. The reason is that they do not only value their material payoffs but they also value the fairness of the behavior of their opponents and the equity of the final outcomes. For example, our theoretical analysis shows that self-interested agents respond much stronger to the implicit incentives provided by a bonus contract. The reason is that – in case that the principal does not reward the agent with a bonus – a fair agent experiences *additional* disutility from the unfairness of the behavior of the principal while the selfish agent “only” suffers from the reduced material payoff. This shows that the presence of fair types 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.

We believe that the effects described in this paper, that have been neglected in contract theory so far, are important for the optimal design of contracts and other governance structures, and that they will prove useful to better understand the provision of incentives in many other contexts as well.



## References

- Aghion, Philippe, Dewatripont, Matthias and Rey, Philippe, 1994. "Renegotiation Design with Unverifiable Information." *Econometrica* 62, 257-282.
- Allen, Franklin and Gale, Douglas, 1992. "Measurement Distortion and Missing Contingencies in Optimal Contracts." *Economic Theory* 2, 1-26.
- Baker, George, Gibbons, Robert and Murphy, Kevin J., 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109, 1125-56.
- Bernheim, B. Douglas and Whinston, Michael D., 1998. "Incomplete Contracts and Strategic Ambiguity." *American Economic Review* 88, 902-32.
- Bolton, Gary E. and Ockenfels, Axel, 2000. A Theory of Equity, Reciprocity and Competition. *American Economic Review* 100, 166-193.
- Camerer, Colin F., 2000. *Behavioral Game Theory*, forthcoming, Princeton, Princeton University Press.
- Charness, Gary, and Rabin, Matthew, 2000. "Social Preferences: Some Simple Tests and a New Model." *mimeo*, University of California at Berkeley.
- Che, Yeon-Koo and Hausch, Donald B., 1999. "Cooperative Investments and the Value of Contracting." *American Economic Review* 89(1), 125-47.
- Dufwenberg, Martin and Kirchsteiger, Georg, 1998. "A Theory of Sequential Reciprocity." Discussion Paper. CentER, Tilburg University.
- Edlin, Aaron S. and Reichelstein, Stefan, 1996. "Holdups, Standard Breach Remedies, and Optimal Investment." *American Economic Review* 86(3), 478-501.
- Falk, Armin and Fischbacher, Urs, 1998. "A Theory of Reciprocity." Discussion Paper. University of Zürich.
- Fehr, Ernst, Gächter, Simon and Georg Kirchsteiger, 1997. "Reciprocity as a Contract Enforcement Device." *Econometrica* 65(4), 833-860.
- Fehr, Ernst and Falk, Armin, 1999. "Wage Rigidity in a Competitive Incomplete Contract Market", *Journal of Political Economy* 107(1), 106-134.

- Fehr, Ernst and Schmidt, Klaus M., 1999. "A Theory of Fairness, Competition and Co-operation." *Quarterly Journal of Economics* 114, 817-868.
- Fehr, Ernst and Schmidt, Klaus M., 2000. "Theories of Fairness and Reciprocity: Evidence and Economic Applications", forthcoming in: M. Dewatripont, L.P. Hansen, S. Turnovski, *Advances in Economic Theory, Eighth World Congress of the Econometric Society*, Cambridge: Cambridge University Press.
- Hart, Oliver and Moore, John, 1999. "Foundations of Incomplete Contracts." *Review of Economic Studies* 66, 115-138.
- Holmström, Bengt, 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13, 324-340.
- Holmström, Bengt and Milgrom, Paul, 1991. "Multi-task Principal-Agent Analyses." *Journal of Law, Economics, and Organization* 7 (Sp.), 24-52.
- Laffont, Jean-Jacques and Tirole, Jean, 1993. *A Theory of Regulation and Procurement*. Cambridge (Mass.): MIT-Press.
- MacLeod, W. Bentley and Malcomson, James M., 1989, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57, 447-480.
- Nöldeke, G., Schmidt, K.M., 1995. Option Contracts and Renegotiation: A Solution to the Hold-Up Problem. *Rand Journal of Economics* 26, 163-179.
- Rabin, Matthew, 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5), 1281-1302.
- Rabin, Matthew, 2000. "Risk Aversion, Diminishing Marginal Utility, and Expected-Utility Theory: A Calibration Theorem", *Econometrica* 68(5), 1281-1292.
- Schmidt, Klaus M. and Schnitzer, Monika, 1995. "The Interaction of Explicit and Implicit Contracts". *Economics Letters* 48, 193-99.
- Segal, Ilya, 1999. "Complexity and Renegotiation: A Foundation for Incomplete Contracts." *Review of Economic Studies* 66(1), 57-82.
- Spier, Kathryn, 1992. "Incomplete Contracts in a Model with Adverse Selection and Exogenous Costs of Enforcement." *RAND Journal of Economics* 23, 432-443.