MANAGERIAL AND DECISION ECONOMICS
Manage. Decis. Econ. **19**: 327–341 (1998)

An Experimental Study of a Dynamic Principal–Agent Relationship

Werner Güth*, Wolfgang Klose, Manfred Königstein and Joachim Schwalbach

Humboldt-University Berlin, Department of Economics and Business Administration, Berlin, Germany

The principal–agent problem is often illustrated by the relationship between owners and managers in modern corporations. Our experimental study considers the problem where the owner has to motivate the manager by an employment contract serving the owner's interest. The contract specifies a salary and a share of the firm's profit which depends on the manager's effort and stochastic market events. The owner can only infer the firm's profitability from the dividend payments. The owner may terminate or change the incentive contract in later periods. The experiment relies on a parameter specification for which risk neutral participants would cooperate efficiently, i.e. the owner should design a contract rendering full effort as optimal. However, we observe contracts for which full effort is not optimal and effort choices which are not optimal. We observe trust and reciprocity as important features of behavior which evolve over a multiperiod principal–agent relationship, but do not carry over to the next game with a new partner.

INTRODUCTION

The standard example of a principal–agent problem is the separation of ownership and control in modern corporations. A large literature on managerial capitalism has been concerned with that issue.[1] It is argued that managers in corporations with highly dispersed ownership have more discretion in pursuing their own interest even if it conflicts with that of the owners. In this situation, incentive contracts can be designed by the owners to motivate managers appropriately.[2] Empirical studies show that financial incentives are the dominant element in management contracts.[3] Although agency theory suggests to relate incentive pay to firm performance or owners' wealth, there is no agreement among empirical studies about the pay-performance relationship. This leads Jensen and Murphy (1990) to the conclusion that '. . . the lack of strong pay-for-performance incentives for CEOs (chief executive officers) indicated by our evidence is puzzling'.[4]

* Correspondence to: Humboldt-University Berlin, Department of Economics and Business Administration, Spandauer Str. 1, D-10178 Berlin, Germany.

Experimental studies may help to solve this puzzle.[5] In our experiment participants are confronted with a dynamic principal–agent situation. It exhibits features of realistic principal–agent problems like multiple periods, stochastic outcomes, imperfect monitoring, contract design, linear reward schemes and downward rigidity of work contracts. Furthermore, the principal decides on how long the partnership lasts and the agent knows more about the state of the enterprise than the principal.

The game has multiple equilibria, but we will nevertheless outline a benchmark solution. We will compare the experimental behavior with this solution, since this gives us some guidance for evaluating results. However, the experiment is not designed as a test of game theory. Our main motivation is to provide empirical facts showing how financial incentives, exclusively considered by most (normative) theoretical studies of principal–agent problems, have to compete with intrinsic motivations, both of the principal and his agent, as well as with social norms. We will see that the players' trust and reciprocity may be even more important than relying on incentive compatibility.

CCC 0143–6570/98/040327-15$17.50

That trust and reciprocity are important behavioral patterns under certain circumstances was also found in a related experimental study by Fehr *et al.* (1997). In this experiment principals simultaneously offer employment contracts, for at most one worker, specifying a wage, a required effort level and a fine which has to be paid if the agent is found (by chance) to be shirking, i.e. not meeting the required effort level. In a stochastically determined order workers can select a contract where there are fewer firms than workers. The worker then chooses his effort level which, finally, is monitored randomly.[6]

In our present study pairs of players interacted for several periods before they were matched with new partners. This allows for reputation formation and in some sense for building up a corporate identity. Furthermore, the principal has almost no direct means to monitor the agent. Thus, our study sheds light on the dynamics of trust and reciprocity within relationships and across relationships, which was not investigated by Fehr *et al.* (1997). The impact of reciprocity on behavior in principal–agent games is also investigated by Fehr and Gächter (1997), who focus on the difference between one-sided and two-sided reciprocity.

Although our experimental game can be varied by changing parameters like the stochastic production function, the cost function of the agent and the time horizon, in this study we varied only the initial value of the firm which is known to the agent but not to the principal. By distinguishing a high and a low value, where the *a priori* probabilities are common knowledge, the agent with the high value has the opportunity to hide low efforts. Furthermore, principal and agent have conflicting interests in the firm's value. While the principal wants to preserve it, since he receives its final value when the game ends, the agent wants to use it for dividend payments, in case he participates in these payments via a profit share.

In the section 'Experimental Design' we describe the experiment in detail. Some behavioral hypotheses are discussed in the section 'Trust and Reciprocity'. The section 'Game Theoretical Analysis' presents the theoretical results for the case in which both parties are guided only by monetary incentives and are risk neutral. In the section 'The Raw Data and Some Stylized Facts' we present the raw data and some stylized facts, and in the section 'Analysis of Effort Decisions' we analyse the agents' effort decisions in detail. The last section contains a summary.

EXPERIMENTAL DESIGN

The participants were divided into two groups of players and seated in two separate lecture halls: rooms A and B. Each B(oss) player took the role of the principal within a principal–agent game and was paired with an A(gent) player. After being seated participants received a letter of welcome which also assigned them their code number. Then we distributed written instructions[7] which informed both players about the rules of the game and that they would interact for at most six periods $t = 1, 2, \dots, 6$.

In period 1 player B, the principal, has to decide about the linear payment scheme (F_1, v_1), the work contract he is offering to player A, the agent. The component F_1 (in general, F_t) is a fixed wage which A receives in period 1 (period t). It does not depend at all on what happens in t. The component v_1 (v_t) is a profit share (a share by which player B participates in the dividend payment P_t in t). In $t = 1$ player B must offer at least one contract (F_1, v_1) with $0 \leq F_1 \leq 4$ and

$$v_1 \in \left\{0, \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}\right\}$$

and can offer at most three contracts of which one has to be accepted by player A. In later periods B can increase F_t and/or v_t, but he cannot reduce either one. Thus, in all periods (F_t, v_t) must satisfy the following restrictions:

$$0 \leq F_t \leq 4 \wedge F_t \geq F_{t-1}$$

and

$$v_t \in \left\{0, \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}\right\} \wedge v_t \geq v_{t-1}.$$

When A decides which work contract to accept he does not know yet the initial value of the firm (W_1). However, it is commonly known that W_1 is low ($W_1 = 3$) or high ($W_1 = 12$) with probability 1/2 respectively.[8] After A's choice among the offered work contracts, he (but not player B) gets to know the actual value of W_1.

A now chooses his effort level $a_t \in \{0, 1, 2, 3\}$ which causes cost

$$K(a_t) = \frac{a_t}{2}$$

on A's side.[9] Then the current market state α_t is determined randomly—with $\alpha_t \in \{1, 2, 3\}$ and probability $\{\alpha_t\} = 1/3$, respectively. The result of the chance move is only disclosed to A. Player A's chosen effort together with the stochastic market state determine the firm's return E_t with $E_t = a_t \alpha_t$. The firm's profit in t is $G_t = E_t - F_t$. Finally, A decides about the dividend payment (P_t). In all periods P_t must satisfy the following restrictions:

$$0 \leq P_t \leq W_t + G_t \text{ if } W_t + G_t > 0$$

and

$$P_t = 0 \text{ otherwise,}$$

where W_t is the firm value at the beginning of t with W_1 determined as above and

$$W_t = W_{t-1} + G_{t-1} - P_{t-1} \quad \forall t > 1.$$

Thus, the agent is free to choose anything between no dividend payment at all or, at the other extreme, running down the firm value all at once. The firm value can actually become negative since G_t can be negative in the case of a positive fixed wage. But, nonetheless the game continues in such a case. More importantly, notice that player B neither observes the initial firm value, A's effort nor the market state. He only observes P_t and in general this allows only imperfect inferences about W_t as well as a_t.

At the end of each period both players can compute their period t incomes which are $\Pi_t^A = F_t + v_t P_t - K(a_t)$ for player A and $\Pi_t^B = (1 - v_t)P_t$ for player B. At the beginning of the next period $(t + 1)$ player B decides whether or not to increase the fixed wage or the profit share or both. Either way player A has to take (F_{t+1}, v_{t+1}) as it is. He chooses his effort a_{t+1} and the game proceeds in the manner described above. Finally, there is a termination rule which allows player B, at the start of periods 4, 5, and 6, to end the game. If he doesn't terminate the game it ends after period 6, i.e. at the beginning of period 7.

Player A's total income is the sum of his period incomes. Player B gets in addition to the sum of his period incomes the end value (liquidation value) of the firm. This is W_7 if the game ends after period 6 or W_4, W_5, or W_6 otherwise. This completes the description of the game.

Subjects played the game twice against changing opponents, but keeping their role. We had 64 participants who were students of economics or business administration (undergraduates and graduates). Thus, we collected 32 plays (one for each principal–agent pair) on each trial. Four subjects, two principals and two agents, formed a matching group, i.e. in the repetition they were cross-matched within their respective group. Thus, the 32 observations of the second trial are not independent. Part of the statistical analysis below will therefore be based on the 16 matching groups.

All monetary values were denoted in so-called 'ECU' (Experimental Currency Unit). The value of one ECU was 0.50 DM in game 1 and 1.00 DM in game 2. The experiment took about 3 h. Average earnings per subject were about 24.50 DM, about US$16 at the time of the experiment (June 1995) for the two games and included a 5.00 DM show-up fee. In the case of efficient game playing the expected average earnings were 48.50 DM per subject, which is also the outcome according to the game theoretic solution (see below). While mean payments per hour were relatively low, they were substantially different between subjects ranging from about 3.10 DM up to 48.60 DM. Given this and the potential gains in case of efficient (equilibrium) play, we consider payment as salient.

After reading the instructions to the game, subjects had to fill out a short questionnaire[10] which asked for a few personal characteristics (such as nationality, sex, and student year) and, more importantly, contained some control questions to check whether subjects understood the rules of the game.

TRUST AND RECIPROCITY

In economics one traditionally assumes rationality and material objectives of the interacting parties. In our experiment all material aspects are captured by monetary incentives. Some implications of rationality based on these monetary incentives will be derived in the next section. Before doing so we want to develop two alternative motives that may guide subjects' behavior: 'trust' and 'reciprocity'. Both may generate behavior that is inconsistent with the implications of rationality and purely monetary motivation.

As demonstrated rather generally by ultimatum bargaining experiments (see, for a survey, Roth, 1995), and more specifically by studies of employment relationship (see Fehr *et al.*, 1997), human decision makers are inclined to respond in kind. They will invest considerable amounts of money in order to punish behavior that may be termed 'greedy', or they reward favorable acts.[11] For the experiment at hand this could occur when a generous contract, offered by the principal, is rewarded by the agent via a considerable work effort.

Responding in kind has been phrased 'reciprocity' (see Young, 1986, for an early reference). Reciprocity as such means that someone's behavior is reciprocated later on by somebody else. So, reciprocity typically relies on sequential interactive decision processes.[12] Usually there is a party moving second which can be said 'to reciprocate' by responding in kind. On the other hand, the party moving first can be said 'to trust' in the other's reciprocity, when choosing an action that makes sense only when it is reciprocated. An illustrative example for trust and reciprocity is the experiment by Berg *et al.* (1995), which, however, offers extremely high incentives for cooperation (the invested amount is tripled).

In our experiment trust is related to the principal offering a fixed wage. More specifically, a higher fixed wage is considered a stronger signal of trust. One might argue that even offering a high profit share could be termed trust. However, we think that the most salient way to express trust is to offer a high fixed wage and no profit share. Furthermore, we do not yet propose trust and reciprocity as sharp theoretical constructs that produce unique predictions for a given game. Rather we propose them as useful labels for observable behavior that carry some intuition. Also, we like to contrast these motives and the associated behavior with the implications of rationality and pure monetary motivation, so that our interpretation seems useful.

In the econometric analysis below, reciprocity will be measured by the responsiveness of the agent's work effort to the size of the fixed wage. We argue that on the basis of rationality and purely monetary incentives there should be no significant influence, while this is called for by reciprocity.

Trust and reciprocity pose some interesting questions for principal–agent theory.[13] Specifically, an essential problem is whether these two motives can be more important for determining behavior than marginal monetary incentives on which normative principal–agent theory is based (see e.g. Hart and Holmström, 1987).

GAME THEORETICAL ANALYSIS

The game is a dynamic game with incomplete information. It exhibits hidden information and hidden actions and it has multiple equilibria. In principle we could apply refinements[14] of the equilibrium concept or the theory of equilibrium selection[15] to narrow down the set of equilibria or even propose a unique strategy profile as the solution to this game. But, looking at the data (see below) we found a great diversity of plays. Thus, any unique prediction will necessarily fail. Instead we will discuss reasonable considerations for the players' choices and characterize a simple equilibrium solution, assuming risk neutral players who are only interested in their own monetary wealth. We will refer to this as the benchmark solution of the game. We will also describe several quite different game paths which are consistent with equilibrium behavior to indicate the manifold of equilibria in this game. Although we will later compare observable behavior with the solution outlined here, the experiment is not designed as a test of game theory, but rather to develop a positive theory of behavior in principal–agent relationships based on experimentally observed data.

Let us first investigate which dividend payment P_t in the interval $0 \leq P_t \leq W_t + G_t$ a rational agent (player A) will select. Clearly, A has no interest in preserving or even increasing the future value W_{t+1}, at least for $t = 6$. He should therefore choose $P_6 = W_6 + G_6$ if this is positive and $P_6 = 0$ otherwise. In earlier periods $t < 6$ a non-maximal P_t could be justified since it might cause improved future contracts (F_t, v_t) or since one may expect a higher dividend share in the future. This, however, does not happen according to the game theoretic solution. Thus, along the solution path

$$P_t^* = W_t + G_t \quad \forall t = 1, \ldots, 6. \tag{1}$$

When the agent determines his effort a_t it seems obvious that his choice depends on the contract (F_t, v_t) he faces. Since marginal cost of effort are 1/2 and the expected productivity is 2, a risk neutral agent chooses $a_t = 3$ if $v_t \geq 3/10$ and $a_t = 0$ if $v_t \leq 2/10$. Thus, given F_t and v_t the agent's best reply behavior $\hat{a}_t(F_t, v_t)$ can be characterized as

$$\hat{a}_t(F_t, v_t) = \begin{cases} 3 & \text{for } v_t = \dfrac{3}{10}, \dfrac{4}{10}. \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Given (2) the principal can induce maximal effort by an appropriate choice of v_t. Therefore in the following we will refer to $v_t \geq 0.3$ as an incentive compatibility constraint. Nevertheless, an agent might choose $a_t = 3$ although $v_t \leq 2/10$, if he expects $v_\tau \geq 3/10$ for some period $\tau > t$. In other words, $v_t = 3/10$ is sufficient, but not necessary for incentive compatibility in the long run. Notice that $\hat{a}_t$ does not vary in the fixed wage component F_t. Paying a fixed wage will not increase effort in the short run. Moreover, incentives can eventually be reduced in the long run. This is possible since a fixed wage reduces the firm value W_t. Over time W_t may become negative such that dividend payments and thereby the agent's work incentives are reduced.

We will show below that along the solution path

$$v_t^* = \frac{3}{10}$$

holds and thus

$$a_t^* = 3 \quad \forall t = 1, \ldots, 6. \tag{3}$$

Regarding the principal's design of the work contract (F_t, v_t) suppose the solution path is

$$F_t^* = 0 \quad \forall t = 1, \ldots, 6 \tag{4}$$

$$v_t^* = \frac{3}{10} \quad \forall t = 1, \ldots, 6. \tag{5}$$

Paying a positive fixed wage or raising the profit share would cause additional cost on the principal's side, but would not increase the agent's effort. Reducing v_t results in zero effort according to (2). Given that (2) holds, if the principal chooses $v_t < 3/10$ at all, he should set $v_t = 0$. Shares $v_t = 2/10$ or $v_t = 1/10$ are clearly suboptimal. Thus, it needs to be checked only whether $v_t = 0$ could increase the principal's profit. Additionally, it remains to be shown that the principal is not interested in early termination of the game along the solution path.

Before we do this we will specify a complete strategy profile. Notice that (1), (3), (4) and (5) together with

'No early termination' (6)

determine a unique path of play which will be the path of play of our proposed solution. To show that this is an equilibrium path we need to specify supporting behavior at unreached information sets.

Suppose that at all unreached information sets (1), (2) and (6) hold as well as

$$F_t^* = F_{t-1} \quad \forall t = 2, \ldots, 6 \tag{7}$$

and

$$v_t^* = \max\left\{\frac{3}{10}, v_{t-1}\right\} \quad \forall t = 2, \ldots, 6. \tag{8}$$

To check that the proposed strategy profile is indeed an equilibrium one might simply work again through our discussion of (1) to (5). It remains to be shown that $v_t = 0$ and early termination are unprofitable for the principal. If $v_t = 3/10$ the agent chooses $a_t = 3$ thereby generating an expected profit in period t of 6 ECU. Thus, the principal gets

$$(1 - v_t)E\{\alpha_t\}a_t = \frac{42}{10} \text{ ECU} \tag{9}$$

each period such that $v_t = 0$ would reduce his profit. Moreover this implies that he should not terminate the game. Termination is possible in periods 4 to 6. In such a case the principal receives the current firm value W_t which is zero at the time according to Equation (1), and gives up positive future profits given by Equation (9).

We will refer to this equilibrium as the benchmark solution. It has a stationary equilibrium path, incentive compatibility is satisfied in each period and maximal effort is chosen always. Thus, the solution is quite simple and efficient. Notice further that according to the solution, dividend payments may be shifted to earlier or later periods without affecting players' payoffs. To render the solution behavior more stable one could rely on marginal impatience assuring that dividends will be paid as early as possible.

There are other equilibria, e.g. equilibria with less than maximal dividend payments and $v_t < 3/10$ in early periods. Suppose $P_t = 0$ and $v_t < 3/10$ in the first two periods and other than that everything is the same as in the benchmark solution. In this case it is still optimal for the agent to choose maximal effort in the first two periods. He simply carries the profits over to the third period and pays them out then. It is also possible to construct equilibria in which the principal pays a fixed wage in addition to a profit share.

In general, it is not obvious which equilibria or equilibrium paths are eliminated by certain refinements. The game has rather many information sets and taking care of all contingencies is tedious. However, one can easily see that paying a fixed wage is excluded by the notion of perfect equilibrium.[16] In every completely mixed strategy profile the behavior at every information set potentially matters. The agent's choice of work effort cannot rationally be conditioned on receiving a fixed wage. Thus, any equilibrium of the perturbed game prescribes not to pay a fixed wage, which establishes the result. Put differently, paying a fixed wage can only be sequentially rational at information sets that are reached with zero probability.

Another aspect one might take into account are the players' attitudes towards risk. So far we assumed risk neutrality for both players. Allowing for different risk attitudes gives more leeway for behavior that can be regarded consistent with game theoretic equlibrium concepts. In the experiment, of course, we could have applied the binary lottery technique, i.e. paying subjects in lottery tickets, in order to implement risk neutrality. However, we did not want to overburden our participants by introducing another stochastic event. Besides there is evidence that the binary lottery technique may be ineffective in inducing optimal decision behavior.[17]

THE RAW DATA AND SOME STYLIZED FACTS

The experimental game underlying our study constitutes a rather rich decision environment. It generated the following data:

- principal's decisions
  - design of initial contract offers
  - contract changes
  - termination of the game
- agent's decisions
  - selection among initial contract offers
  - work efforts
  - dividend payments.

We deliberately chose a complicated game to study behavior within a laboratory environment that exhibits features of realistic principal–agent problems like multiple periods, stochastic outcomes, imperfect monitoring, contract design and downward rigidity of work contracts. These features make the relation between work contracts and work efforts as studied by principal–agent theory less obvious for experimental subjects. While we think it is important to study experimental behavior in complex games this creates some difficulties for the statistical analysis. Within a rich environment many factors might influence behavior. Thus, when we try to discover the influence of one factor we need to control for the influence of others. We do this in detail when applying regression analyses to explain the agents' effort choices.

The analysis of the principals' decisions is less thorough. How he adapts the contract conditional on his experience within the repeated game seems an interesting but complicated task. Especially, here, principals could not reduce any component of the contract[18], and in addition they were allowed to terminate the game in some periods.

Nevertheless, at first, we will briefly describe the raw data for all decisions and characterize some observable features as stylized facts. Most of these facts rely on mean data and simple descriptive statistics, but some are behavioral hypotheses for which we provide test statistics. Thereafter we will turn to the analysis of effort decisions.

Initial Contract Offers

Eighty-one percent of principals in the first game and 75% in the second game offered their agent more than one contract. The number of offers is significantly reduced in game 2 ($p = 0.059$, Wilcoxon Matched-Pairs Signed-Ranks Test, one-tailed). More than 90% of principals offered positive fixed wages, i.e. among the contracts offered by each single principal at least one exhibited a positive fixed wage. The average offered fixed wage is significantly lower in game 2 ($p = 0.023$, Wilcoxon Matched-Pairs Signed-Ranks Test, one-tailed).

All principals, with one exception in game 2, offered a positive profit share. Game 2 contracts were not significantly different from game 1 decisions. However, the fraction of offered contracts that satisfy incentive compatibility as implied by Equation (2) is smaller in game 2 than in game 1. These observations are summarized as:

Table 1. Means (S.D.) of Profit Share and Fixed Wage by Game and Period

Period	Game 1		Game 2	
	Profit share	Fixed wage	Profit share	Fixed wage
1	0.21 (0.12)	1.42 (1.06)	0.19 (0.12)	1.07 (0.72)
2	0.23 (0.12)	1.43 (1.05)	0.22 (0.09)	1.28 (0.80)
3	0.26 (0.10)	1.54 (1.10)	0.26 (0.09)	1.42 (0.85)
4	0.29 (0.10)	1.63 (1.09)	0.32 (0.08)	1.51 (0.84)
5	0.32 (0.08)	1.73 (1.02)	0.32 (0.08)	1.57 (0.87)
6	0.31 (0.08)	1.91 (0.89)	0.34 (0.06)	1.71 (1.01)

Stylized Fact 1 *Contrary to the theoretical analysis, initially offered contracts exhibit substantial fixed wages and non-incentive compatible profit shares. This holds even for experienced subjects, but experience reduces the number of offers and the average fixed wage.*

Choices among Initial Offers and Self-Selection

Among those agents who received more than one contract offer (F_1, v_1) 19% decided for the contract with the highest profit share. All others, given they had a choice, went for the highest fixed wage. One can interpret this observation as self-selection by agents. Agents who intend to choose low effort levels obviously prefer a higher F_1. For this reason one should expect agents who gave up some fixed wage for a higher profit share to choose higher effort levels. This is actually the case in period 1 ($p = 0.006$, Mann-Whitney U Test, one-tailed). We did not check this for later periods since the effect of self-selection could be confounded with other influences.

Stylized Fact 2 *Subjects' preferences for work contracts are heterogenous and seem related to their intended work effort. In part, this is signalled via self-selection.*

Fixed Wages and Profit Shares

Table 1 displays means for fixed wages and profit shares for each game and period separately. Profit shares and fixed wages increase over time which is not surprising since by the rules of the game they are not allowed to decrease.

Fixed wages are quite high contrary to $F_t^* = 0$. Mean profit shares are below the incentive compatible level $v_t^* = 3/10$ in the beginning and slightly above this level in the end. These observations lead to:

Stylized Fact 3 *In general, work contracts differ from the benchmark solution. Many principals pay positive and increasing fixed wages. While some profit shares are below the incentive compatible level others exceed it. This holds even for experienced players.*

Effort Levels

Figure 1 shows a bar chart of the chosen effort levels in both games. If agents would have be-

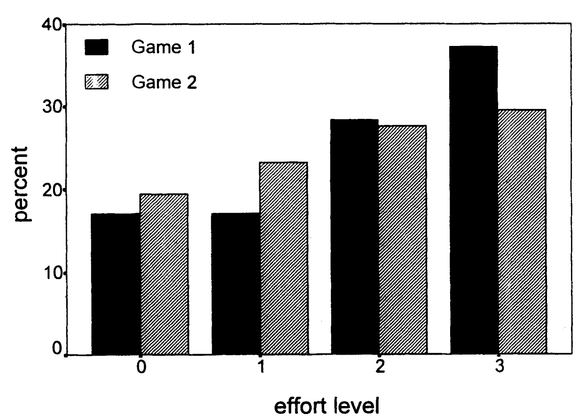

Figure 1. Distribution of the agents' effort decisions.

Table 2. Means (S.D.) of Dividend Payments and Firm Values by Game, Period and Initial Firm Values

Period	Low initial firm value ($W_1 = 3$)		High initial firm value ($W_1 = 12$)	
	Dividend	Firm value	Dividend	Firm value
Game 1				
1	1.80 (1.61)	3.22 (2.33)	3.67 (3.77)	10.15 (3.98)
2	1.68 (1.56)	3.39 (3.13)	4.17 (2.89)	8.24 (3.94)
3	2.97 (2.52)	3.51 (4.16)	5.23 (3.89)	5.35 (4.01)
4	2.56 (3.52)	2.70 (4.95)	4.03 (2.29)	3.92 (4.63)
5	2.26 (2.47)	1.36 (5.91)	5.07 (5.93)	0.33 (0.96)
6	2.22 (2.03)	1.27 (5.23)	0.89 (1.24)	−0.70 (0.97)
Game 2				
1	1.71 (1.89)	3.03 (1.80)	4.36 (4.62)	10.39 (4.60)
2	1.31 (1.76)	3.39 (3.17)	4.09 (5.04)	8.29 (6.63)
3	1.72 (1.99)	2.65 (4.17)	3.53 (4.30)	6.76 (7.71)
4	1.15 (0.95)	1.89 (3.58)	2.84 (3.30)	5.78 (8.56)
5	1.62 (2.93)	0.99 (3.85)	5.00 (5.64)	1.87 (3.53)
6	1.62 (3.01)	−1.17 (1.47)	4.35 (5.30)	−0.20 (0.76)

haved optimally, we should observe either zero or maximal efforts and nothing in between. Instead, we see a lot of intermediate choices.

In game 1 (game 2) 19 (16) out of 32 agents chose $a_t = 1$ or $a_t = 2$. This supports:

Stylized Fact 4 *A considerable fraction of effort decisions are inconsistent with short run incentive compatibility as implied by Equation* (2).

Dividend Payments and Firm Values

Table 2 displays means for dividend payments and firm values separately for each game, period and initial firm value. The average firm value decreases over time. At the end of game 2 (repetition) it is negative while it is still positive at the end of game 1 (first trial) meaning that some agents gave up money in favor of their principals. That experimental subjects hesitate to fully exploit their strategic advantage is known from other studies (see the survey of Roth, 1995). However, looking closer at the data we found that in game 1 (game 2) out of 15 (14) agents who reached period 6 only 2 (1) left a positive firm value. All others brought it down to zero or had already run it down earlier.

If the initial firm value is high the agent might signal this—at least imperfectly—to the principal by choosing rather high dividend payments.[19] But instead average dividend payments are moderate and a more detailed look at the individual data reveals that signalling in this sense hardly occurred. These observations support:

Stylized Fact 5 *Contrary to the benchmark solution, dividend payments are less than maximal each period. Agents do not signal a high firm value. They rather seem to avoid this.*

Termination of Games

Only about half of the plays lasted six periods. All others were terminated earlier by the principals. Table 3 presents the relative frequencies for termination, conditional on reaching periods 4 to 6. It exhibits a quite sharp increase at the end of both games. Intuitively one might conjecture that the principal's decision to terminate the game depends somehow on the dividend payments, since these are the agent's actions which he can observe. To see whether zero dividends result more likely in termination than positive dividend we computed these conditional probabilities.

In game 1 the probability for termination in t after $P_t = 0$ was 23%, while it was 18% after $P_t > 0$. In game 2 the probability for termination in t after $P_t = 0$ was 41%, while it was only 12% after $P_t > 0$. This supports:

Table 3. Termination of Games[a]

Period	Game 1 (%)	Game 2 (%)
4	3.1	18.8
5	12.9	11.5
6	44.4	39.1

[a] Rel. frequency conditional on reaching period t.

Stylized Fact 6 *Principals regarded zero dividend payments as a negative signal and thus, more likely terminated the game after $P_t = 0$ than after $P_t > 0$.*

ANALYSIS OF EFFORT DECISIONS

As indicated above in a more detailed analysis we want to address the following questions:

1. *What determines the agent's effort or, conversely, induces more shirking?*
2. *What is the relation between effort and profit share as well as between effort and fixed wage?*

From a normative point of view contracts should be incentive compatible in order to induce maximal efforts. More specifically, the agent's choice of work effort would solely depend on his profit share. However, relying on empirical evidence the principal might pay a fixed wage and might hope that the agent will honor this trust, i.e. reciprocate the principal's favorable action by choosing a high effort level. Thus, we explore whether experimental behavior relies more on incentive compatibility or more on trust in reciprocity and the corresponding reaction by agents.

That observed behavior does not literally correspond to the normative predictions became already clear by looking at the raw data. We observed positive fixed wages, non-incentive-compatible profit shares and intermediate effort decisions. Another simple test of principal–agent theory is to check whether chosen efforts were zero when the profit share was below 0.3 and whether efforts were maximal otherwise. In game 1 (game 2) conditional on receiving a profit share below 0.3 only 17% (24%) of agents chose zero effort. Conditional on a higher profit share only 44% (39%) chose maximal effort. Taken together about two thirds of all effort decisions in both games are inconsistent with incentive compatibility (and risk neutrality). However, these observations could have been generated by relying on trust and reciprocity.

Facing the great diversity of subjects' behavior no simple explanation seems adequate, and thus, we ran regression analyses to control for the influence of several factors.

Data Aggregation

Running regressions bears some problems here since each agent decided upon his effort level up to six times within one game. Since the subjects played two games within the session this means that we collected up to 12 effort decisions of a single agent. In statistic terms this is 'repeated measurement' and in regression analysis it might result in correlated error terms. Another problem is that four subjects—two principals and two agents—formed a matching group such that their strategic interaction could result in correlated error terms as well. To deal with the problem of strategic interaction we aggregated the data of all individuals within each matching group for each half of the two games. For all variables we computed separately for each matching group the means for:

- periods 1 to 3 in game 1, i.e. 'phase 1 data',
- periods 4 to 6 in game 1, i.e. 'phase 2 data',
- periods 1 to 3 in game 2, i.e. 'phase 3 data' and
- periods 4 to 6 in game 2, i.e. 'phase 4 data'.

Thus, instead of 328 individual decisions (32 agents times 12 decisions minus 56 missing observations due to early termination) we analyse 16 time series of length 4, i.e. 64 observations. Each time series is comprised of the mean data of one matching group. Thus, it is independent of each other time series which implies that there is no problem of strategic interaction anymore. What is left is the problem of repeated measurement and this has to be taken care of by an appropriate econometric method for statistical analyses that combine time series and cross-sectional data.

Model 1

We chose the following fixed-effect specification:

$$a_{it} = \beta_0 + \delta_1 v_{it} + \delta_2 F_{it} + \sum_{j=1}^{4} \gamma_j D_{jit}^{\text{time}} + \sum_{j=1}^{2} \omega_j D_{jit}^{\text{treat}} + \sum_{j=1}^{16} \mu_j D_{jit}^{\text{group}} + e_{it} \qquad \text{(Model 1)}$$

with the variables:

a_{it} = effort of matching group i in phase t
v_{it} = profit share of matching group i in phase t
F_{it} = fixed wage of matching group i in phase t
D_{jit}^{time} = dummy variables for phases 1 to 4 with

$$D_{jit}^{\text{treat}} = \begin{cases} 1 & \text{if } j = t \\ 0 & \text{otherwise} \end{cases}$$

Table 4. Regression Results of Model 1

Variable		Coefficient	Standard error	p value
Constant		0.32	0.43	0.233
Low initial firm value		−0.45	0.18	0.008
Fixed wage		0.58	0.13	0.000
Profit share		2.56	1.22	0.021
Time effects				
Constant	$t=2$	−0.41	0.18	0.013
	$t=3$	0.02	0.16	0.443
	$t=4$	−0.45	0.20	0.015
Individual fixed effects				
Matching Group 1		—	—	—
—		—	—	—
Matching Group 16		—	—	—
Summary of fit				
R^2: 0.75			adj. R^2: 0.62	
F-ratio: 5.85			Probability > F: 0.000	
Observations: 64				

D_{jit}^{treat} = dummy variables for treatments 1 and 2 (high or low initial firm value) with

$$D_{jit}^{\text{treat}} = \begin{cases} 1 & \text{if matching group } i \text{ in treatment } j \\ 0 & \text{otherwise} \end{cases}$$

D_{jit}^{group} = dummy variables for matching groups 1 to 16 with

$$D_{jit}^{\text{group}} = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

e_{it} = error term

and where the coefficients measure: β_0 = global constant term; δ_1 = influence of profit share; δ_2 = influence of fixed wage; γ_j = fixed effect for phase j; ω_j = fixed effect for treatment j; μ_j = fixed effect for matching group j.

Model 1 cannot be estimated in the current version, since it is not identified due to linear dependence among the dummy variables and the unit vector (which is implicitly contained in the model to estimate the global constant term). We follow the usual procedure of omitting enough variables to get identification. The regression result is displayed in Table 4.[20] The model is highly significant and explains a considerable part of the variance.

The agent's choice of work effort increases with the profit share as well as with the fixed wage. Both correlations are positive and significant. Thus, increasing the fixed wage by 1.00 DM, increases effort by 0.58 units (on average). Accordingly, a 0.1 increase in profit share rises effort by about 0.26 units (on average).

To measure the influence of time we chose dummy variables instead of a single time variable. This allows for more general time effects and indeed the influence of time is not monotonic. Effort decreased from phase 1, the reference category, to phase 2 and from phase 3 to phase 4, but not from phase 2 to phase 3. Effort decreased in the course of playing a game, but not across games. Reasons for this time pattern could be a termination effect or some kind of learning or experience effect. That efforts are higher in the beginning of a game than in the end can be explained by an agent's initial, but false expectation that the profit share rises during the game.

Agents who received a low initial firm value, $W_1 = 3$, chose lower efforts[21] than agents receiving a high initial firm value, $W_1 = 12$. There seem to be two reasonable explanations for this fact. A simple one is purely psychological—namely that getting a low firm value causes frustration and lethargy. A second and more subtle economic reason is that a low initial firm value is run down faster than a high one with the consequence that $W_t - F_t$ becomes negative. If this is the case, part of the profit in t cannot be used for dividend payments (see the section 'Trust and Reciprocity'). Actually, if the deficit is very large, profits cannot be used for dividend payments at all. Thus, the agent's expected gains from working, i.e. his work incentives, are reduced. Of course,

Table 5. Regression Results of Model 2

Variable		Coefficient	Standard error	p value
Constant		1.22	0.83	0.075
Low initial firm value		−1.40	0.18	0.016
Time effects				
Constant	$t=2$	−0.81	1.26	0.264
	$t=3$	−0.79	0.86	0.184
	$t=4$	−2.57	1.12	0.014
Profit share	$t=1$	0.12	2.13	0.475
	$t=2$	0.21	2.56	0.465
	$t=3$	2.93	2.31	0.106
	$t=4$	6.41	2.05	0.002
Fixed wage	$t=1$	0.34	0.26	0.095
	$t=2$	0.68	0.25	0.005
	$t=3$	0.42	0.35	0.116
	$t=4$	0.54	0.18	0.003
Individual fixed effects				
Matching Group 1		—	—	—
—		—	—	—
Matching Group 16		—	—	—
Summary of fit				
R^2: 0.80			adj. R^2: 0.64	
F-ratio: 5.23			Probability > F: 0.000	
Observations: 64				

according to the benchmark case this will never occur. But if it happens (and it did in the experiment) a rational agent would reduce his effort.

The dummy variables for the matching groups were included to account for repeated measurement. Their joint inclusion improves the coefficient of determination R^2 by 25%, and the improvement is highly significant according to a partial F-Test ($p = 0.003$). This shows that it was important to take care of repeated measurement. We do not want to interpret these variables further than this and therefore refrained from reporting their coefficients and test statistics (six coefficients are statistically significant).

Of course, we tried to include other regressors like the market state. But they showed no significant influence.

Overall the regression model explains the data reasonably well. The influences of profit share and fixed wage are both positive, i.e. they are qualitatively of the type one would expect from a normative respectively behavioristic perspective. However, it was not clear whether both will prove to be substantial and statistically significant. The influence of time indicates dynamic aspects of behavior that we want to discuss in more detail.

Model 2

It is quite plausible that behavior changes slightly, due to learning during the course of the experiment. Learning could be different when responding to the profit share than when reacting to the fixed wage. Thus, it appears questionable whether phase t constants are adequate to measure learning effects. As an alternative we estimated the following model:

$$a_{it} = \beta_0 + \sum_{j=1}^{4} \delta_{1j} D_{jit}^{\text{time}} v_{it} + \sum_{j=1}^{4} \delta_{2j} D_{jit}^{\text{time}} F_{it} + \sum_{j=1}^{4} \gamma_j D_{jit}^{\text{time}} + \sum_{j=1}^{2} \omega_j D_{jit}^{\text{treat}} + \sum_{j=1}^{16} \mu_j D_{jit}^{\text{group}} + e_{it} \qquad \text{(Model 2)}$$

with all variables and the coefficients β_0, γ_j, ω_j and μ_j as in model 1 and where the coefficients δ_{1j} and δ_{2j} measure, respectively, the influence of profit share in phase j and the influence of fixed wage in phase j.

The regression result is displayed in Table 5.[22] The model is highly significant and explains a considerable part of the variance. The difference from model 1 is that now the influence of fixed wage and profit share are allowed to vary across time. Notice, that, for example, the variable 'profit share in $t = 3$', which is displayed in Table 5, corresponds to the term $D_{jit}^{\text{time}} v_{it}$, for $j = 3$, in the above equation (model 2). Thus, the variable 'profit share in $t = 3$' exhibits the actual profit share value for phase 3 observations and is coded as zero for observations other than those of phase 3.[23]

In modelling the influence of time (via fixed wages and profit shares) we chose again a dummy variable specification to allow for more general time effects. While in model 1 both the profit share as well as the fixed wage exhibited statistically significant influences, model 2 splits these influences into phasewise effects resulting in a number of insignificant regressors. However, looking at the coefficients and their t-values shows striking regularities—qualitatively as well as quantitatively. First of all, the influences are qualitatively the same as expected and already confirmed by model 1. The influences of profit share and fixed wage are positive in all phases while a low initial firm value affects work effort negatively.

The importance of the profit share is monotonically increasing over time. While it is negligible in size and statistically insignificant especially in the first two phases (in game 1) it is substantial and significant in the end. The phase 4 coefficient of 6.41 implies that an increase of 0.1 in the profit share increases the agent's effort by 0.64 units on average. Taking together the result of regression 1 and the regular pattern of the profit share influence we draw the following conclusion:

Conclusion 1 *The agent's choice of work effort is positively influenced by work contracts which offer higher profit shares. This correlation gets stronger over time indicating that agents might eventually learn incentive compatibility over time.*

The influence of fixed wages is substantial and significant from the start—although it is insignificant in phase 3. From the perspective of the principal this means that paying higher fixed wages is a reliable way to induce higher efforts which needs no learning. The fixed wage influence increases, but not in each phase. It decreases from the end of game 1 to the beginning of game 2 (from phase 2 to phase 3). This suggests that the agent's responses to fixed wages have a dynamic aspect which is relationship-specific. The principals' trust, signalled via a positive fixed wage, is reciprocated by the agent via high efforts. This positive reciprocity is stronger in the long run than in the short run, but it is not carried over from one relationship to the next. It has to be built up afresh when one confronts a new partner. One might argue that there is some carry-over effect since the influence of fixed wages is slightly stronger at the beginning of game 2 than in the beginning of game 1, but one should not interpret our regression results too boldly. After all not all of the coefficients are significantly different from zero. Also, the differences between the coefficients are mostly not significant[24] (what can easily be checked by comparing the differences and the standard errors of the coefficients).

Taking together the results of regression 1 and the regularities of the fixed wage influence we further conclude:

Conclusion 2 *The agent's choice of work effort is positively influenced by work contracts which offer higher fixed wages. This correlation is already strong at the start of the game.*

Conclusion 3 *Trust (measured by the size of the fixed wage) and reciprocity (measured by the responsiveness of work effort on the size of the fixed wage) exhibit dynamic aspects which are relationship-specific. Reciprocity gets stronger in the course of a game, but is not carried over across games (with a new partner).*

By now, and with regard to the first question posed above, we found a reasonably good explanation for the agents' effort decisions. Regarding the second question we found that both higher profit shares as well as higher fixed wages increase agents' efforts.

- But does either or both increase the principal's profits?
- Will the principal prefer to rely on incentive compatibility than on trust and reciprocity?

More specifically let us examine whether it would be more profitable for the principal to increase the fixed wage rather than the profit share given that the agent's reaction is as suggested by our regression results and assuming average behavior. Let $\bar{v}_t$ and $\bar{P}_t$ represent means of profit share and dividend payment in phase t.

Thus, the principal's average income $\bar{\Pi}_t^{\mathrm{B}}$ (not including the liquidation value) is given by

$$\bar{\Pi}_t^{\mathrm{B}} \equiv (1 - \bar{v}_t)\bar{P}_t$$

Following regression model 2, changes in fixed wage ΔF_t or profit share Δv_t result in changes of effort Δa_t and accordingly in changes of return ΔE_t and the firm's profit ΔG_t. Compared to $\bar{\Pi}_t^{\mathrm{B}}$ the principal's income induced by a change of the work contract $\bar{\Pi}_t^{\mathrm{B}}(\Delta)$ can be written as

$$\bar{\Pi}_t^{\mathrm{B}}(\Delta) \equiv (1 - \bar{v}_t - \Delta v_t)(\bar{P}_t + \Delta G_t)$$

with $\Delta G_t = \Delta E_t - \Delta F_t$. A contract change is profitable if $\bar{\Pi}_t^{\mathrm{B}}(\Delta) - \bar{\Pi}_t^{\mathrm{B}} \geq 0$ which is equivalent to

$$(1 - \bar{v}_t - \Delta v_t)\Delta G_t - \Delta v_t \bar{P}_t \geq 0. \tag{10}$$

Now we will check whether it is profitable to induce a small increase in effort of, for example, $\Delta a_t = 0.1$ units leading to an expected increase in return of $\Delta E_t = 0.2$ DM. According to regression model 2 this can be achieved by the following contract changes:

	ΔF_t (DM)		Δv_t (%)
in $t = 1$:	0.292 DM	or	83.3
in $t = 2$:	0.147 DM	or	47.6
in $t = 3$:	0.238 DM	or	3.4
in $t = 4$:	0.185 DM	or	1.6

Furthermore, note that the mean values $\bar{v}_t$ and $\bar{P}_t$ are:

	$\bar{P}_t$ (DM)	$\bar{v}_t$ (%)	$(1-\bar{v}_t)\cdot 0.2/(\bar{P}_t+0.2)$
in $t = 1$:	3.185	23.3	0.045
in $t = 2$:	2.089	30.5	0.042
in $t = 3$:	2.636	22.6	0.055
in $t = 4$:	2.630	32.2	0.048

The values in the right column for the column-head expression are reported for convenience (see below). One now can use these data to evaluate condition (10). For changes in fixed wage this condition reduces to

$$(1 - \bar{v}_t)\Delta G_t \geq 0$$

where

$$\Delta G_t = \Delta E_t - \Delta F_t = 0.2 - \Delta F_t$$

so that the above inequality just requires $0.2 \geq \Delta F_t$. Thus, on average, inducing an increase in effort via the fixed wage would have been profitable in $t = 2$ and $t = 4$, but not in the early phases of each game ($t = 1$ and $t = 3$). For $\Delta v_t > 0$ and $\Delta F_t = 0$ the corresponding condition is

$$\frac{(1 - \bar{v}_t) \cdot 0.2}{\bar{P}_t + 0.2} \geq \Delta v_t.$$

So, to induce a small effort increase via the profit share causes a very substantial loss in game 1 ($t = 1$, $t = 2$), but is profitable in game 2 ($t = 3$, $t = 4$).

Thus, if the agent is inexperienced (game 1), the principal should rather rely on trust and reciprocity than on incentive compatibility. Inducing more effort via a higher profit share reduces his profit substantially. With experience (game 2) a profit share policy seems on average more profitable.

SUMMARY

While principal–agent theory is exclusively built on incentives, actual behavior is sometimes better explained by trust on the side of principals and reciprocity on the side of agents. Incentive compatibility has to be learnt whereas trust and reciprocity is readily available even for inexperienced participants. Within an ongoing relationship trust and reciprocity get more important over time, but the experience of this kind of mutual cooperation is not carried over to another relationship with a new partner. Rather it has to be built up afresh. Expressed differently, the dynamic aspects of trust and reciprocity are stronger within relationships than across. Within our experimental framework neither the process of learning incentive compatibility nor the dynamics of trust and reciprocity can be regarded as settled. This certainly needs further investigation.

After all, we found that laboratory behavior within this quite complex game follows systematic patterns that are based on a few concepts. While some of these findings may seem natural from a psychologic point of view, they appear puzzling from a purely economic perspective (focussing only on monetary incentives). It would certainly be premature to discuss their implications for principal–agent relationships in real life. But nevertheless our evidence indicates that even when the principal can hardly monitor the agent's performance, behavior is not purely opportunistic in an economic sense. Since principals may quite often encounter situations where they must trust their agents, this is surely good news.

Acknowledgements

We thank Simon Gächter, Ulrike Grasshoff, Rolf Tschernig and Jürgen Wolters as well as the editor, Stuart Mestelman, and an anonymous referee for their helpful comments. Furthermore, we thank Vital Anderhub, Steffen Huck, Martin Strobel and Stefan Winter for their support in running the experiment, and also we gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft, Projekt 'Strategisches Handeln' and SFB 373.

NOTES

1. See Marris and Mueller (1980).
2. See Jensen and Meckling (1976) and Fama and Jensen (1983).
3. See Finkelstein and Hambrick (1988) for a literature review.
4. Jensen and Murphy (1990) (p. 252).
5. For other experiments testing principal–agent theory see Berg *et al.* (1992), Epstein (1992), Fehr *et al.* (1993), and Keser and Willinger (1997).
6. To avoid negative monetary incentives of their experimental participants the authors rely, however, on inconsistent rewards, which we consider a weakness of the design. Namely, what the agent receives is not what the principal must pay.
7. All instructions and other information sheets used for running the experiment including a preexperimental questionnaire are available from the authors.
8. The probabilities were implemented by announcing that W_1 will be 3 (respectively 12) for half of the players.
9. The game will proceed this way in each period, thus, we substitute the subscript 1 by t.
10. Available from the authors.
11. Some evidence for this type of behavior can even be found in interaction between animals. See, for instance, de Waal (1982).
12. Sometimes even cooperation in simultaneous move games, e.g. in prisoners' dilemma experiments, is viewed as reciprocity. It is argued that a player may anticipate his opponent to reward own cooperativeness.
13. See Osterloh and Frey (1997) for a related discussion of intrinsic versus extrinsic motivation within firms.
14. See van Damme (1987).
15. See Harsanyi and Selten (1988).
16. See Selten (1975).
17. See Güth *et al.* (1993) as well as Selten *et al.* (1995).
18. In statistical terms this leads to a censored regression model which is substantially complicated by the fact that censoring here depends on previous contract choices.
19. In special cases the agent can even make the principal perfectly aware of a high initial firm value, e.g. by choosing $P_1 > 12$ which is impossible if the initial firm value is low.
20. One might wonder whether linear regression analysis is adequate here. Individual effort is a discrete variable that can take on only four values, which could result in a lot of boundary observations. However, first one should note that we aggregated six individual effort decisions (two player by three periods). Thus, mean effort can take on 19 different values and we observe only six (out of 64) boundary observations. Second, looking at the residuals convinced us that the linear regression model does reasonably well.
21. The problem of how to measure the treatment effect correctly, as it is discussed in Königstein (1997), does not arise here, since the dummies for the matching groups are non-nested within the treatment dummies. For some groups the initial firm value was low in game 1, but high in game 2, and vice versa.
22. Again, we removed enough variables from the model to get identification.
23. As in model 1 we included matching group dummies (results are not reported) to control for repeated measurement. Again, as a group they substantially improved the model fit (by 22%) and the change in R^2 was highly significant ($p = 0.005$, partial F-test, one-tailed).
24. This problem also arises in the discussion of how the influences of profit shares v_t changes over time.

REFERENCES

J.E. Berg, L.A. Daley, J.W. Dickhaut and J. O'Brien (1992). Moral hazard and risk sharing: experimental evidence. In *Research in Experimental Economics* (edited by R.M. Isaac), Greenwich, CN: JAI Press, 5, pp. 1–34.

J.L. Berg, J. Dickhaut and K.A. McCabe (1995). Trust, reciprocity and social history. *Games and Economic Behavior*, **10**, 122–142.

F. de Waal (1982). *Chimpansee Politics*, Baltimore: Johns Hopkins University Press.

S. Epstein (1992). Testing principal–agent theory. In *Research in Experimental Economics* (edited by R.M. Isaac), Greenwich, CN: JAI Press, 5, pp. 35–60.

E.F. Fama and M.C. Jensen (1983). Agency problems and residual claims. *Journal of Law and Economics*, **26**, 327–349.

E. Fehr and S. Gächter (1997). How effective are trust- and reciprocity-based incentives? In *Economics, Values and Organizations* (edited by A. Benner and L. Putterman), Cambridge: Cambridge University Press.

E. Fehr, G. Kirchsteiger and A. Riedl (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, **108**, 437–459.

E. Fehr, S. Gächter and K. Kirchsteiger (1997). Reciprocity as a contract enforcement device: experimental evidence. *Econometrica*, **65**, 833–860.

S. Finkelstein and D. C. Hambrick (1988). Chief executive compensation: a synthesis and reconciliation. *Strategic Management Journal*, **9**, 543–558.

W. Güth, E. van Damme and M. Weber (1993). The normative and behavioral concept of risk aversion–an experimental study. Working Paper, Tilburg University.

J.C. Harsanyi and R. Selten (1988). *A General Theory of Equilibrium Selection in Games*, Cambridge, MA: MIT Press.

O. Hart and B. Holmström (1987). The theory of contracts. In *Advances in Economic Theory*, *Fifth World Congress* (edited by T. Bewley), Cambridge: Cambridge University Press.

M.C. Jensen and W.H. Meckling (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, **3**, 305–360.

M.C. Jensen and K.J. Murphy (1990). Performance pay and top-management incentives. *Journal of Political Economy*, **98**, 225–264.

C. Keser and M. Willinger (1997). Principals' principles when agents' actions are hidden. Working Paper, University of Karlsruhe.

M. Königstein (1997). Measuring treatment-effects in experimental cross-sectional time series. Working Paper, Humboldt-University at Berlin.

R. Marris and D.C. Mueller (1980). The corporation, competition and the invisible hand. *Journal of Economic Literature*, **18**, 32–63.

M. Osterloh and B.S. Frey (1997). Managing motivation: crowding effects in the theory of the firm. Working Paper, University of Zürich.

A.E. Roth (1995). Bargaining experiments. In *The Handbook of Experimental Economics* (edited by J.H. Kagel and A.E. Roth), Princeton, New Jersey: Princeton University Press, pp. 253–348.

R. Selten (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, **4**, 25–55.

R. Selten, A. Sadrieh and K. Abbink (1995). Money does not induce risk-neutral behavior, but binary-lotteries do even worse. Discussion Paper, SFB 303, No. B-343, Bonn University.

E. van Damme (1987). *Stability and Perfection of Nash Equilibria*, Berlin: Springer Verlag.

G.A. Young (1986). Patterns of threat and punishment reciprocity in a conflict setting. *Journal for Personality of Social Psychology*, **51**, 541–546.

