

An Experimental Bribery Game

by KLAUS ABBINK, BERND IRLBUSCH, and ELKE RENNER

October 1999

Abstract

Essential characteristics of corruption are (1) reciprocity relationships between bribers and public officials, (2) negative welfare effects, and (3) high penalties when discovered. We separate the influences of these factors in an experiment. In a two-player game reciprocation is economically inefficient through negative externalities. A control treatment without externalities is also conducted. In a third, so-called *sudden death* treatment, corrupt pairs face a low probability of exclusion from the experiment without payment. The results show that reciprocity can establish bribery relationships, where negative externalities have no apparent effect. The penalty threat significantly reduces corruption, although discovery probabilities are typically underestimated.

Keywords

Corruption, reciprocity, fairness, repeated games

JEL Classification Codes

C91, D62, D72, D73, K42

Acknowledgements

The authors thank Jose Apesteguia, Heike Hennig-Schmidt, Robin Pope, Bettina Rockenbach, Abdolkarim Sadrieh, Reinhard Selten, and seminar participants in Barcelona, Bonn, Cologne, and Grenoble for helpful comments and suggestions. All errors remain our own. Financial support by the *Deutsche Forschungsgemeinschaft* through the *Sonderforschungsbereich 303*, the European Union through the TMR research network ENDEAR (FMRX-CT98-0238), and the *Land Nordrhein-Westfalen* is gratefully acknowledged.

Addresses

Klaus Abbink	Bernd Irlenbusch	Elke Renner
Laboratorium für experimentelle Wirtschaftsforschung Universität Bonn Adenauerallee 24-42, D-53113 Bonn fax: 49-228-73-9193		Lehrstuhl für Organisationstheorie Otto-Beisheim-Hochschule (WHU) Burgplatz 2, D-56179 Vallendar fax: 49-261-6509-111
phone: 49-228-73-9192 abbink@lab.econ1.uni-bonn.de	phone: 49-228-73-9196 bi@uni-bonn.de	phone: 49-261-6509-303 erenner@whu-koblenz.de

“Thou shalt not wrest judgement; thou shalt not respect persons, neither take a gift; for a gift doth blind the eyes of the wise, and pervert the words of the righteous.”

DEUTERONOMY 16. 19.

1. Introduction

Though it is not at all new, the phenomenon of corruption has attracted growing attention during the nineties. The breakdown of communist systems and the end of the cold war have induced several implications which raised increasing awareness of the problem. Donor countries do not turn a blind eye to corrupt practices in receiver governments anymore just because the latter belong to the correct political camp. Liberalisation of media in many countries has made the public much more sensitive and better informed than before. Moreover, due to the globalisation of the economy, politicians around the world have recognised that corruption is no longer an internal problem of the single countries, hence governments showing a *laissez-faire* attitude came under international pressure¹. The growing alertness is reflected by an increasing number of international initiatives to combat bribery. Some of them had already a practical impact: for example, the OECD Anti-Bribery Convention of 1997 ended the common practice to subsidise bribes to foreign officials by tax deductibility. In the 29 countries having signed the convention corresponding laws have been released, with Germany as the latest country to prohibit tax deductibility of bribes.

As corruption is widely recognised as a major economic problem around the world, there is also a growing interest in the empirical analysis of its causes and consequences. In recent years, significant progress has been made by the statistical analysis of field data. MAURO (1995) finds statistical evidence for the growth-reducing effect of corruption. LONG and RAO (1995) expose that the value of firms significantly fall after corrupt behaviour of managers has been discovered. VAN RIJCKEGHEM and WEDER (1997) find corruption to be significantly lower when public officials' salaries are high relative to manufacturing wages. GUPTA, DAVOODI, and ALONSO-TERME (1998) analyse the effect of corruption on the income distribution in a society, and find evidence that corruption increases poverty.

While the analysis of field data is helpful to identify socio-economic determinants and consequences of corruption, it gets extremely difficult if we wish to study corrupt behaviour itself. Naturally, corruption is hard to observe in the field, since everyone engaged in it has good reasons to remain silent. Therefore, we introduce an alternative approach to the empirical analysis of corruption. In an interactive laboratory experiment, we can observe behaviour in simulated bribery scenarios directly. Further, we can identify which variables influence corrupt behaviour by changing the experimental environment. In this paper, we start with examining behaviour

¹ For a discussion of these and other arguments see TANZI (1998).

in a very basic bribery scenario. Consider a firm who applies for a licence to perform some environmentally harmful activity. Or think of a firm selling products to the government. In such cases, the public official has some discretion, where her decision has crucial impact on the profits of the firm. On the other hand, the public official receiving a flat salary has little or no personal benefit if she refuses the permission to pollute the environment, or if she purchases from the cheapest supplier. Thus, by illegal payment of bribes in return to favourable decisions, firm and official can achieve mutual improvement at the cost of the public: In the first case, the population suffers from pollution, in the second case, the society of taxpayers bears the too high prices charged by the bribe-giving firm.

The examples show three essential characteristics of corruption. First, there is the *reciprocity* feature: both firm and official can exchange benefits with one another. However, since bribery is illegal, enforceable contracts on corrupt acts cannot be made. The exchange of favours must rely on trust and reciprocation between briber and bribee. Second, corruption imposes serious *negative externalities* to the public. Third, bribe-giving and -taking is liable to severe penalties in case of discovery. Thus, corruption is inherently *risky*.

In this paper, we design an experiment to separate the influence of these three characteristics on the behaviour of (potential) bribers and public officials. For this purpose, we construct three variants of a simple sequential game. In all three variants, one player (the special interest party) can transfer money to a second player (the public official) in order to induce a decision advantageous to him. This option, however, is costly for the public official. With these features alone, we obtain a very simple reciprocity game, which allows us to study the impact of trust and reciprocity in the corruption situation separately. In a second treatment, the same game is played, with the extension that the option favourable to the special interest party induces a negative externality to other individuals, in a way that the efficiency losses exceed the mutual benefits of the two players directly involved. By comparing this treatment to the basic one, we separate the effect of social welfare considerations on the behaviour of individuals. Finally, to investigate the effect of the danger of being caught, we introduce a new feature which we call the *sudden death* treatment. When a bribe has been paid and accepted, a lottery is played out, and with very low probability the activity is discovered. In that case, both players' accounts are cleared, and they are excluded from any further play. By comparison to the other treatments, we can examine the effects of a penalty that is unlikely, but drastic, on corruption behaviour. All three treatments were played in thirty-round supergames between the same players. Thus, we model a long-term relationship between a briber and a public official.

To our knowledge, experimental corruption games have not emerged in the economic literature.² Most economic studies on corruption involve merely theoretical analyses of situations

² A so-called *quasi-experiment* has been conducted by LÜDTKE and SCHWEITZER (1993). In a questionnaire study, the authors elicit the subjects' tolerance towards corrupt activities. A (non-interactive) individual decision experiment has been conducted by FRANK and SCHULZE (1998) to examine differences in the corruptibility of economists and students from other disciplines.

conducive to corruption.³ The behaviour of the agents is given by the assumption of profit or utility maximisation. This assumption is useful and often plausible, but sometimes empirically problematic. Especially since corruption involves a reciprocity relationship, we have to consider the experimental evidence against self-interested rationality in such situations. The tradition of experimental reciprocity games starts off with a study by FEHR, KIRCHSTEIGER, and RIEDL (1993), who analyse reciprocal fairness in a labour market context. Firms can pay a wage to workers, who then can specify their effort level arbitrarily. The authors observe that experimental firms systematically overpay workers, compared to the competitive equilibrium wage. Workers reciprocate by an increased effort. As a result, mutual improvement of firm and worker is achieved at the price of underemployment.

In the *investment game* by BERG, DICKHAUT, and MCCABE (1995) the first moving player can pass money to the second mover, which is tripled by the experimenter. The second mover can voluntarily return money to reward trust. The results clearly refute the hypothesis of subgame perfect rationality and support the impact of reciprocal fairness. Other experimental studies on similar games have followed these seminal papers⁴. All examining different questions, their common finding is a strong impact of trust and reciprocation even in one-shot situations, in which completely anonymous players meet for playing the game only once. Repeated interaction in a reciprocity game, as we analyse in the present study, is examined in a recent study by MEIDINGER, ROBIN, and RUFFIEUX (1998). The subjects play a simplified version of the original investment game in a 25-round supergame. The authors find that trust and reciprocation are effective in the beginning of the experiment, but their effect decreases with time.

These studies already suggest that although bribery is built on trust in absence of self-enforcing devices, reciprocity may establish stable bribery relationships. However, reciprocity is only one aspect of corruption, therefore these findings cannot be transferred directly to corruption scenarios. In the existing experimental reciprocity games, reciprocal exchange is not only mutually beneficial, but also increases economic efficiency. This is fundamentally different in the exchange of benefits between briber and official. Further, corruption involves an exogenous risk through the possibility of getting caught, which is not modelled in typical reciprocity games. Thus, a new experiment has to be designed.

2. The Experimental Model

As the simplest bribery situation, we consider a public official who has only two options (e.g. accepting or rejecting a proposal), where it is in her discretion alone which one to implement.

³ Standard references are ROSE-ACKERMAN (1975, 1978) or SHLEIFER and VISHNY (1993). All models in the theoretical literature known to us are essentially different from ours. The model most akin is probably the one by MANION (1997) who analyses a one-to-one relationship between a briber and a public official.

⁴ see e.g. JACOBSEN and SADRIEH (1996), FEHR, GÄCHTER, and KIRCHSTEIGER (1997), HOFFMAN, MCCABE, and SMITH (1998), FAHR and IRLBUSCH (forthcoming), ABBINK, IRLBUSCH, and RENNER (forthcoming), or DUFWENBERG and GNEEZY (forthcoming).

A potential briber (a user of public services, typically a firm, for simplicity we will speak of “users” or “firms” in the following) can transfer money to the official in the hope that this would influence her decision to his advantage. We conducted the experiment in three treatments: the *pure reciprocity* (PR) treatment serves as a control treatment. Second, the *negative externality* (NE) treatment damages to other players are introduced, to separate the effect of social welfare considerations. In the *sudden death* (SD) treatment, finally, an external risk is introduced. In all treatments, this situation is played 30 times subsequently.

2.1. The Pure Reciprocity Treatment

At the first stage, the firm decides on whether or not to transfer an amount t of money to the public official, and if yes, how much he wishes to transfer. If he transfers a positive amount, he must pay a small “transfer fee” of 2. The fee represents the initiation costs of the briber when he approaches the civil servant to establish a reciprocal relationship. These costs can be considered as being independent from the later course of play, i.e. they must be paid also if the official should reject the bribe.

The public official is then asked whether she accepts or rejects the transfer. If she rejects, the transfer is not performed, both accounts remain unchanged apart from the firm paying the small transfer fee. If the official accepts, then she receives the tripled amount. The factor reflects a difference in marginal utility: the same amount of money means much less to a relatively larger firm than to a public official⁵. To keep things simple, only integers could be transferred. The maximum amount to be transferred also had to be restricted to ensure that the firm could not get negative cumulative payoffs by transferring too much. Only transfers out of the interval [0...9] were feasible.

At the second stage, the public official has to make a binary decision between two alternatives, simply called X and Y. Y is much more favourable to the firm, X is slightly better for the official. This constellation reflects that the briber’s advantage from Y is immense, but the public official has only a slight preference for X arising from some effort necessary to justify a Y choice before her superiors. In the parameter constellation of the experiment, X induces utilities of 36 for both players, Y leads to utilities of 56 for the user and 30 for the official.

2.2. The Negative Externality Treatment

In this treatment, the second feature of a bribery scenario is introduced. Whereas in the pure reciprocity treatment it is overall efficient to reciprocate (and hurts nobody else), we now introduce a negative externality to the public arising from reciprocal behaviour. We do this by simply imposing a fixed monetary damage on each of the other subjects in the session whenever

⁵ In the later treatments, the multiplier also ensures that no negative total payoffs can occur.

the public official chooses Y. All other participants of the experiments therefore play the role of the public for each single pair⁶.

As the laboratory has 18 seats, there are nine pairs of subjects playing the game simultaneously. Thus, if a pair chooses Y, 16 other subjects are damaged. The damage was set to 3, such that a choice of Y induced a total damage of 48. A single subject can be damaged by at maximum 24 per round, if in all eight other pairs in the session the public official chooses Y. By the damages, the Y alternative is – in terms of total payoff for all players – overall inefficient, since the total damages exceed the mutual gains realised by the two players of a pair.

As bribery is carried out secretly, we provided no feedback about the other players' decisions. No subject was ever informed about whether other subjects had been corrupt, thus the subjects did not know to what extent they were damaged by other pairs. Consequently, they were not informed about their exact current cumulative earnings during the play.

2.3. The Sudden Death Treatment

In the third treatment of our experiment, the danger of being caught at corrupt activities is modelled. To keep things simple, we model discovery as an exogenous lottery. The question arises how such a lottery should be designed. We decided for a new feature which we call the *sudden death* treatment, using an analogy from sports⁷. Whenever a transfer of a positive amount is accepted (irrespective of whether the civil servant's decision has been influenced by it), a lottery is played out determining whether the corruption activity is discovered. In case of discovery (which happens with very low probability), both the user and the public official are disqualified from the experiment. This means that not only both players lose all their earnings from the experiment, but they are moreover excluded from further play. Thus, they also lose any chance to re-gain the money they lost in later rounds. The sudden death penalty reflects features that seem important to us for an appropriate modelling of the real situation. First, complete disqualification is supposed to be the most drastic penalty that is doable in an experiment, given there are legal and practical restrictions to impose negative total payoffs to subjects. Second, the sudden death treatment captures that real life agents face the end of their current careers in case of discovery.

The probability for the sudden death event, 2 , was set to $2 = 0.003$ in each round in which a positive transfer was accepted. By this, it was ensured that only very risk-averse subjects would refrain from being corrupt because of their fear of being disqualified. Denote by r the number of rounds in which a positive offer is accepted. The probability of falling prey to the

⁶ This feature emphasises the “public bad” character of corruption. Everybody – including the public official and the firm themselves – would be better off if everybody were not corrupt, but for any given behaviour of everybody else, it is profitable for the single pair of firm and official to exchange benefits at the costs of others. For a survey on public bad and public good experiments see LEDYARD (1995).

⁷ In the overtime of a tied ice hockey match, the team scoring the first goal wins the match. This rule is known as the *sudden death* rule because the losing team has no further chance to cure the loss.

sudden death is $1 - (1 - 2)^t$. At maximum, for $r = 30$, we obtain an overall sudden death probability of $p = 0.086$ for $2 = 0.003$. Denote by x the total payoff received by equilibrium play, and by y the additional payoff that a player can gain if the pair plays a reciprocation strategy. A risk-neutral individual would prefer reciprocation in every round if $(x + y)(1 - p) > x$, hence $y/x > p/(1 - p)$. For $p = 0.086$, a reciprocity gain of 9.4% would suffice to make reciprocation more attractive than equilibrium play, given it could be established. As an example, by a transfer of 6 talers every round and a following Y choice, both players would realise a gain between 33.3% and 100%, depending on the behaviour of the other pairs.

Figure 1 depicts the game tree of the sudden death treatment stage game. “U” is the user, “P” the public official. “C” denotes a chance move. The hangman pictogram illustrates the event of sudden death. The lines “-3 ... -3” mean that all other subjects are damaged by 3 talers. Note that the game tree is incomplete in the sense that the payoffs shown at the end nodes depict only the payoff consequences induced by the two players’ moves. In addition, the final payoffs are influenced also by the decisions of all other pairs, which is not noted in the tree.

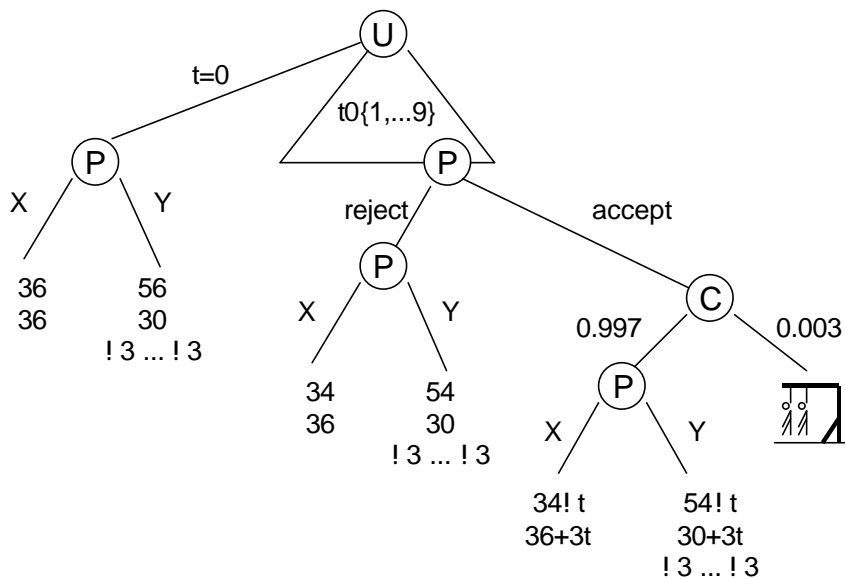


Figure 1. The (incomplete) game tree of the sudden death treatment

2.4. Game Theoretic Analysis

It is straightforward to see that it is an equilibrium if all firms always transfer zero and all public officials always choose X. If all other players follow this behavioural pattern, then any unilateral deviation of a the user by transferring a positive amount would only reduce her payoff by at least the transfer fee (no matter whether the public official’s strategy involves accepting or rejecting the transfer), as well as any official who would choose Y would forego a payoff of 6 without any possible gain.

However, it is also easy to show that on *any* equilibrium path no positive transfers and no Y choices are made, as the following proposition states.

Proposition. *In an equilibrium of the bribery supergame, the user never transfers a positive amount, and the public official never chooses Y in information sets that are reached with positive probability.*

Equilibria exist which involve positive transfers or Y choices at some information sets, but these information sets then cannot be reached on an equilibrium path. Therefore, the only outcome that can be rationalised is the non-corrupt one. Note that this holds without applying any refinement to the Nash equilibrium concept. A detailed proof is given in appendix A.

Through the unambiguous game theoretic property of the equilibrium path the game draws an idealised picture of a bribery situation in which corrupt relationships cannot be established by rational profit maximisation alone. Rather, they require trust and reciprocation. Therefore, the game allows us to study the impact of reciprocity in a corrupt relationship, since corruption out of trust and reciprocation is sharply contrasted to equilibrium behaviour.

3. Experimental Procedure

The experiment was conducted in November 1998 at the *Laboratorium für experimentelle Wirtschaftsforschung* at the University of Bonn. The subjects were encouraged to participate through posters in the campus. Most of them were students from various disciplines, where law and economics students constituted the largest fractions.

The experiment was computerised with software developed using *RatImage* (ABBINK and SADRIEH 1995). The game was presented in a way that all possible moves of both players of one pair were visible on one screen. Yellow triangles marked the way through the experiment. After a decision was made, the stages that could not be reached anymore were low-lighted. After all decisions of a round had been made, the subjects were informed about their payoffs resulting from their own pair's decisions, and (in the treatments with negative externalities) they were reminded that their payoffs would also be influenced by the decisions of all other pairs in the experiment. A screenshot of the display used for the sudden death treatment is reproduced in appendix B.

To minimise presentation effects, we designed our experiment in a completely context free fashion. We presented the corruption situation to the experimental subjects without connecting it to any story. The actors were just called player 1 and player 2, without mentioning the roles they play as public official and special interest party, and without explicitly calling a transfer of money from player 1 to player 2 a "bribe". We decided for a neutral presentation to avoid the

uncontrolled effects of possible connotations raised by hypothetical stories, and to ensure best possible comparability with other experimental results⁸.

Each session began with an introductory talk. The written instructions and payoff tables that listed the payoffs for all possible strategy combinations were handed out to the subjects. Translations of both are provided in appendix B (original text in German). The instructions were read aloud and explained in detail. After the introduction, the subjects were seated in cubicles, visually separated from one another by curtains. The terminal numbers, which determined the role of that subject as being user or public official, were attributed to the subjects by random draw. After the subjects were seated, the play started immediately. The thirty rounds of the experiment were played in slightly less than an hour, such that a whole session took about 1½ hours including introduction.

In the sudden death treatment, subjects could possibly be disqualified from the experiment. To ensure that those subjects who were disqualified did not leave the session (which would have been noticed by the other subjects and thus distorted the statistical independence of the pairs), a lump sum show up fee of DM 5.- was paid to every subject. The show up fee was provided mainly to give the disqualified subjects an incentive to remain seated until the end of the session. The disqualified subjects had to fill in an on-screen questionnaire during the time they were waiting. The questionnaire comprised a number of questions concerning the subject's attitude towards fairness, taken from KAHNEMAN, KNETSCH, and THALER (1986a and 1986b), and tasks of logical inference copied from internet experiments by psychologists from the University of Bonn. The questionnaires were meant to keep disqualified subjects busy rather than to collect meaningful data.

Immediately after the session, the subjects were paid anonymously in cash. The exchange rate was set to 0.025 DM per taler in the pure reciprocity treatment, and 0.03 DM per taler in the other two treatments, for which, through possible damages, lower final payoffs had to be expected. The total earnings in the session ranged from DM 5.00 (one pair of subjects was unlucky in the sudden death treatment) to DM 47.77 with an average of DM 33.22 for 1½ hours, which is considerably more than a student's regular per hour wage in Bonn. One DM is equivalent to EUR 0.51.

Two sessions with 18 subjects were conducted in each of the three treatments. Since no feedback from other pairs is given, every single pair can be considered as a statistically independent observation. Thus, we gathered 18 observations in each of the three treatments.

⁸ It is known that the presentation of a game can have a significant impact on subjects' behaviour. A famous early example is provided by PRUITT (1970), who finds that the presentation of the prisoners' dilemma game has a significant impact on the likelihood of co-operation. For a recent study on the effects of wording in the experimental instructions see BURNHAM, McCABE, and SMITH (1998). Interestingly, framing seems to play a role not only in experiments, but also in real life corruption cases. It is supposed to be important to avoid that an approach is perceived as bribery (NECKEL 1995) This is often done by giving bribes in form of non-monetary gifts (this issue is discussed in OFFER 1997 and; for a practical "guide" see DAVIS 1997).

4. Hypotheses

The three treatments of our experiment allow to test several hypotheses by comparison of the treatments. In addition, we have requested the subjects to fill in a questionnaire to test a hypothesis that was put forward in the literature.

The first hypothesis addresses the reciprocity aspect of the corruption situation. The literature on reciprocity games suggests that exchange of benefits is observable also if it violates individual rationality. We might hypothesise that this force is strong enough to establish corrupt relationships, despite the absence of enforcement devices.

Hypothesis 1. *Firms tend to transfer substantial positive amounts, which is reciprocated by the public officials' choice of Y. This effect is particularly strong in the PR treatment.*

It is well-known that individuals' behaviour is often affected by social considerations. Thus, we might expect that the damage done to others in the negative externality treatment will reduce the frequency of Y choices, and, as a reaction, also lower the transfers made by the users. We formulate our second hypothesis accordingly.

Hypothesis 2. *In the negative externality treatment, public officials tend to choose Y less frequently, and users tend to transfer less than in the pure reciprocity treatment.*

Although the sudden death probability was chosen very low, we might expect that its very existence induces a corruption-detering effect, since the penalty is drastic. We therefore formulate the third hypothesis as follows.

Hypothesis 3. *In the sudden death treatment, public officials tend to choose Y less frequently, and users tend to transfer less than in the negative externality treatment.*

KIRCHGÄSSNER (1997) hypothesises that in real-life corruption cases, individuals might systematically under-estimate the probability of being caught. He argues that the frequent occurrence of corruption cases in reality is most likely due to a wrong calibration of the discovery probability, since the negative consequences are so severe that they can hardly be compensated by the additional income realised by bribe-taking, if the possibility of discovery would be sufficiently taken into account.

It may be conjectured that in their calibration of the discovery probability, individuals are tempted to anchor to the discovery probability in a *single* case, rather than considering the total probability of being caught at some point during the long-term relationship. In our experimental framework, such an effect could be due to an inappropriate consideration of the fact that the event of sudden death can happen in *every* round in which a bribe is accepted.

To test this hypothesis, we requested the subjects to estimate the overall probability of being caught for different parameter constellations. After the play of the sudden death treatment, all subjects who were not disqualified had to fill in a questionnaire with nine questions of the following kind: "Suppose player 1 always transfers a positive amount, and player 2 always accepts. How large do you estimate the probability that these players are disqualified during the

whole session, if the experiment lasts n rounds and the probability of disqualification is $p\%$ in each round?"

We asked nine questions with the nine possible combinations of $n \in \{10, 50, 100\}$ and $p \in \{0.1\%, 1\%, 5\%\}$. If the answer was not farther away from the correct answer than 10 per cent points, the subject was rewarded with 10 talers, for an exactly right answer 20 talers were awarded. To exclude effects arising from the order in which the tasks are presented, we randomised the order for each subject independently.

Following KIRCHGÄSSNER's conjecture, we formulate the last hypothesis.

Hypothesis 4. *The subjects tend to under-estimate the total probabilities of being disqualified in a whole session, for given single-round probabilities and round numbers.*

Because answering the questionnaires was incentivised with money, the disqualified subjects did not do the probability estimation tasks. This would have spoilt the character of the sudden death as preventing them from any further chance to earn money in the session.

5. Results

The level of corruption manifests itself in the amounts of bribes that are offered and in the frequency of (manipulating) Y choices induced by bribes. In the following, we will analyse the experimental results with respect to the four hypotheses formulated in the previous section. The raw data are available from the authors upon request.

5.1. The Impact of Reciprocity

Hypothesis 1 states that corrupt relationships can be established through trust and reciprocity, even if there is no means to enforce reciprocation. The tables 1 and 2 show that this hypothesis is clearly confirmed. Table 1 shows the average transfers per round made by the single firms. Table 2 shows the average frequency of Y choices by the single public officials. All numbers are ordered from the lowest to the highest value in a treatment. The strong impact of reciprocity can be identified through the strong correlation between average transfer offer and average Y choice frequency across the pairs. The higher the transfer in a pair, the higher tends to be the frequency of Y choices. The Spearman rank correlation coefficients are $r_s = 0.87$ in the PR treatment, $r_s = 0.94$ under the NE condition, and $r_s = 0.93$ in the SD treatment. All coefficients are significantly positive at less than $\alpha = 0.01$. Figure 2 depicts the distribution of the amounts transferred by the firms, for the aggregate data of all pairs in the single treatments, figure 3 depicts the relative frequency of Y choices after a certain transfer has been offered. In the NE and the PR treatment, the modal transfer is 6 talers⁹, which is a clear deviation from the equilibrium

⁹ The predominance of this offer value can be explained by the fact that with a transfer of 6 both players' final payoffs are equal, given the public official accepts and reciprocates by choosing Y. For a discussion of the equal payoff principle see SELTEN (1978).

prediction, according to which no corruption would occur. Also in the SD treatment, deviations from equilibrium behaviour are observed in more than half of the rounds.¹⁰ It can clearly be seen that public officials tend to reciprocate by choosing Y after they received relatively high transfers, while they typically choose X after they received no or small bribe offers.

Table 1. Average transfer offer by individual users, over all rounds

PR		NE		SD	
Pair	Avg. transfer	Pair	Avg. transfer	Pair	Avg. transfer
3	0.0	5	0.3	1	0.0
12	1.0	3	1.6	7	0.0
14	1.8	9	1.9	16	0.3
11	3.3	11	3.1	8	0.5
18	3.5	14	3.5	9	0.7
6	3.9	15	3.5	18	1.3
2	4.2	17	4.6	12	1.6
17	4.2	16	4.8	13	2.1
15	4.7	6	5.2	2	3.2
16	5.0	10	5.2	11	4.1
4	5.2	4	5.3	15	4.2
7	5.8	7	5.3	17	4.4
10	5.8	13	5.4	5	4.5
5	6.0	8	5.5	3	4.6
9	6.2	1	5.6	4	5.0
1	6.7	12	5.8	14	5.0
8	6.9	2	6.0	10	5.4
13	8.2	18	8.7	6	5.8
Avg. st.dev	4.58 2.12	Avg. st.dev	4.52 1.94	Avg. st.dev	2.93 2.08

Table 2. Frequency of Y choices by individual public officials, over all rounds

PR		NE		SD	
Pair	Avg. Y freq.	Pair	Avg. Y freq.	Pair	Avg. Y freq.
3	0.00	5	0.13	1	0.00
12	0.03	3	0.17	16	0.00
14	0.20	9	0.20	7	0.03
18	0.43	15	0.40	8	0.03
6	0.47	17	0.40	9	0.07
11	0.47	11	0.47	12	0.07
7	0.67	14	0.50	13	0.23
17	0.67	6	0.73	18	0.33
2	0.70	4	0.80	2	0.47
15	0.73	8	0.80	15	0.50
9	0.77	16	0.80	3	0.70
13	0.87	10	0.83	5	0.70
16	0.87	7	0.87	10	0.70
1	0.90	13	0.87	17	0.70
4	0.90	1	0.90	11	0.73
10	0.93	18	0.90	4	0.77
5	1.00	2	0.97	14	0.83
8	1.00	12	0.97	6	0.93
Avg. st.dev	0.65 0.32	Avg. st.dev	0.65 0.29	Avg. st.dev	0.43 0.34

In the 30 round supergame, not only the public official can reciprocate on the user's bribe offers, but also the firm can condition his transfer decision on the observed choice made by the official in previous rounds. If such user reciprocation was present in the data, we should observe higher transfers as a reward for previous Y choices, and low transfers as a punishment for experienced X choices. For every single user, we measure the impact of his level of reciprocation as the difference between the average transfer after a Y decision and the average amount sent after a X choice by the public official in the preceding round. Formally, our measure of user reciprocation is computed as

$$R = \frac{\sum t_Y}{\#Y} - \frac{\sum t_X}{\#X}$$

where t_Y denotes the transfer after a preceding Y choice, t_X is the transfer after a X choice, $\#Y$ and $\#X$ denote the number of Y and X choices in rounds 1-29. A measure for R can only be

¹⁰ The equilibrium outcome ($t=0$; X) is observed in 44.8% of the rounds.

computed if a user experiences at least one Y and one X choice during the rounds 1 to 29. A high impact of user reciprocation would show up in positive values of the R measure.

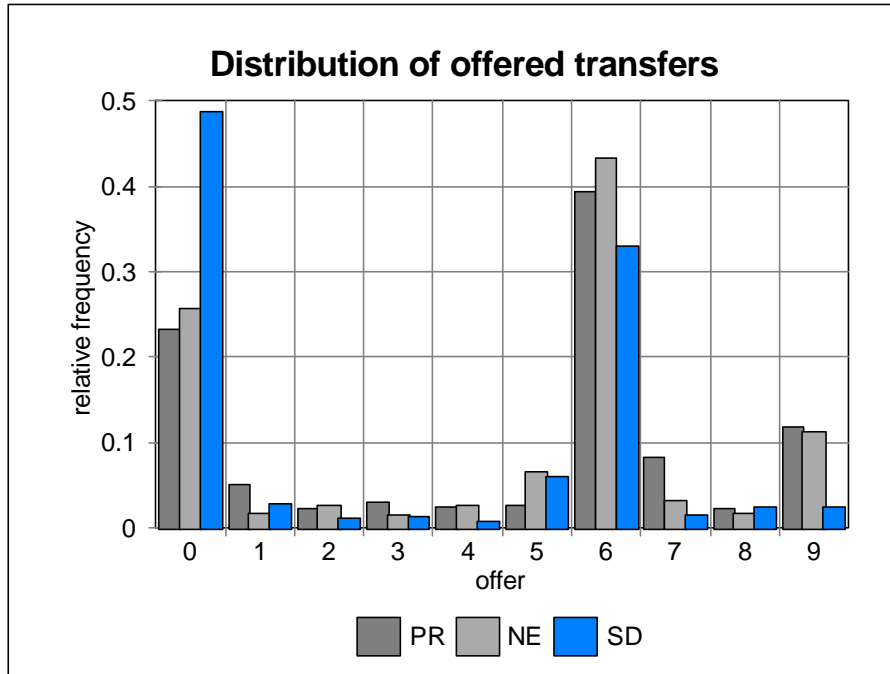


Figure 2

On average, we obtain R measures of +2.16 (PR treatment), +3.30 (NE), and +3.43 (SD). In all treatments, the average R value is clearly positive. The Wilcoxon matched pairs signed rank test applied to the R values of the individual firms rejects the null hypothesis of equal probability of positive and negative measure at a significance level of $\alpha = 0.001$ (one-sided) in all three treatments.

Interestingly, there seems to be a treatment difference between the PR condition and the two treatments with negative externality, the latter exhibiting a higher impact of user reciprocation. In fact, the Mann-Whitney U-test applied to the individual R measures rejects the null hypothesis of equally high reciprocation measures in favour of the alternative hypothesis of lower measures in the PR treatment at a one-sided significance level of $\alpha = 0.10$ (PR versus NE) and $\alpha = 0.05$ (PR versus SD). It seems that the presence of a negative externality causes more pronounced user reciprocation. Without negative externality, users tend to condition their transferring behaviour less on their experience in the preceding round. The damage done to others seems to strengthen the sensitivity for kind and unkind behaviour between the two players in one pair. However, we do not find an entirely convincing explanation for this phenomenon, especially since it seems to be the case only for user behaviour.

Observation 1. *Reciprocity establishes stable corrupt relationships. In all treatments, public officials reciprocate on higher bribes by choosing Y, users reciprocate on favourable choices by paying higher bribes in the next round.*

5.2. The Impact of Negative Externalities

According to hypothesis 2, subjects' considerations of social welfare should lead to lower corruption levels in the NE than in the PR treatment. The figures 2 and 3 already suggest that this is not evidently the case. The distribution of bribes being offered as well as the responses by public officials on them are similar in the PR and the NE treatment. On average, 4.58 talers are transferred in the PR treatment, compared to 4.52 talers under the NE condition. The public officials choose Y in 64.5% of all rounds in the PR and 65.0% in the NE condition. Of course, the differences are not significant. Thus, we do not observe an apparent influence of the negative externalities on decision behaviour.



Figure 3

The figures 4 and 5, which show the development of corruption levels in terms of average bribe offers (figure 4) and average frequency of Y choices (figure 5) provide an even clearer picture. Not only the overall level of corruption, but also its development over time is remarkably similar in both treatments. Both graphs show a tendency of rising corruption levels in the beginning of the experiment, which are stable after approximately ten rounds¹¹ until an apparent end

¹¹ If we compute the change in average bribe offers and average Y choice frequencies from the first to the second block of ten rounds for the single pairs, we obtain positive changes for the majority of users in the PR

effect forces both bribe offers and Y choice frequencies downwards in the very last rounds of play¹². If we correlate the average transfer in round i in the PR treatment to the corresponding average transfer in the same round in the NE treatment, we obtain a Spearman rank correlation coefficient of $r_s = 0.59$, which is significant at $\alpha = 0.01$ (one-sided).¹³

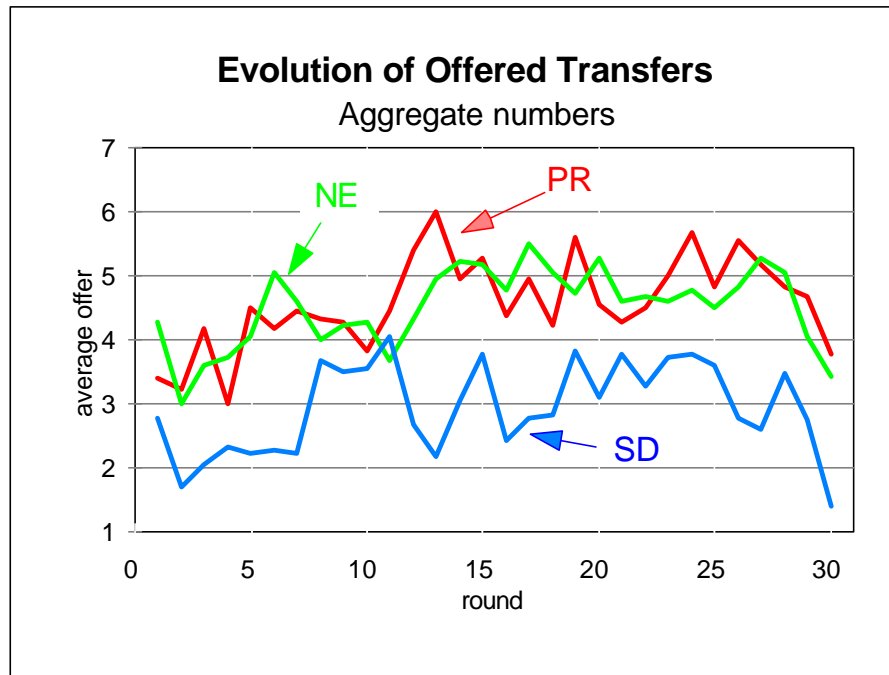


Figure 4

Observation 2. *The negative externality on other subjects has virtually no effect on the level of corruption. The distribution of bribe offers and the frequency of Y choices are very similar in the treatments PR and NE, both in absolute terms and in their evolution over time.*

5.3. The Impact of Penalties

Hypothesis 3 states that we should observe lower levels of corruption in the sudden death treatment because subjects fear the penalty. In fact, this hypothesis is supported by the data. On average, 2.93 talers are offered under the sudden death condition, roughly 35% less than in each of the two treatments without sudden death. The average frequency of Y choices falls by about the same rate down to 43.8%. The Mann-Whitney U-test rejects the null hypothesis of equal transfers (Y choice frequencies, resp.) at a significance level of at least $\alpha = 0.05$ (one-sided)

and NE treatment. The Wilcoxon matched pairs signed rank test rejects the null hypothesis of equal probability of positive and negative signs at a significance level of $\alpha = 0.05$ for both treatments and both measures.

¹² An end-game effect is characterised by a breakdown of co-operation towards the last rounds of the experiment. For a discussion see SELTEN and STOECKER (1986).

¹³ A weak, but significant correlation ($r_s = 0.33$, significant at $\alpha = 0.05$, one-tailed) can also be detected between the PR and the SD treatment. The correlation between the treatments NE and SD is not significant.

for both measures and both pairwise comparisons of the SD and one of the other treatments. Thus, the fear of the sudden death significantly reduces the tendency to pay bribes.

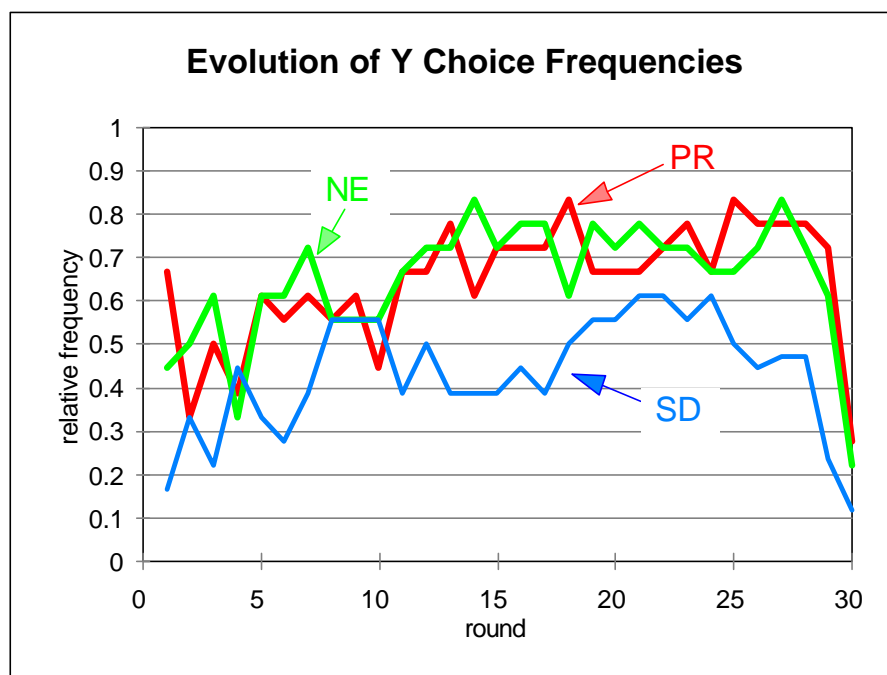


Figure 5

The fear of getting penalised is also expressed in another aspect of our data. Figure 6 suggests that public officials in the SD treatment exhibit a higher tendency to reject bribe offers than those in the other conditions. The absolute number of rejections is significantly higher than both in the PR treatment ($\chi^2 = 0.10$) and the NE treatment ($\chi^2 = 0.05$, one-tailed, according to the Mann-Whitney U-test). More specifically, we can see that higher rejection frequencies can also be observed after relatively high transfer offers, which are most plausibly rejected to avoid the sudden death lottery.¹⁴ Weakly significantly more transfers of at least 6 talers are rejected in the SD than in both the PR and the NE treatment ($\chi^2 = 0.10$, one-tailed, according to Fisher's exact test applied to the number of pairs in which an offer of at least 6 talers was rejected at least once). Thus, the threat of the sudden death penalty induces a higher tendency to reject bribes.

Observation 3. *The threat of the drastic penalty significantly reduces the level of corruption. The average bribe offer and the frequency of Y choices are decreased by the penalty. Public officials exhibit a higher tendency to reject bribes.*

¹⁴ Small bribe offers might also be rejected to signal that the official considers the amount offered by the firm as too low (if the official rejects the transfer, the user cannot be tempted to misunderstand the X choice as trust exploitation). This would explain why small offers are frequently rejected also in the PR and NE treatment, where rejecting bribes means leaving money on the table.

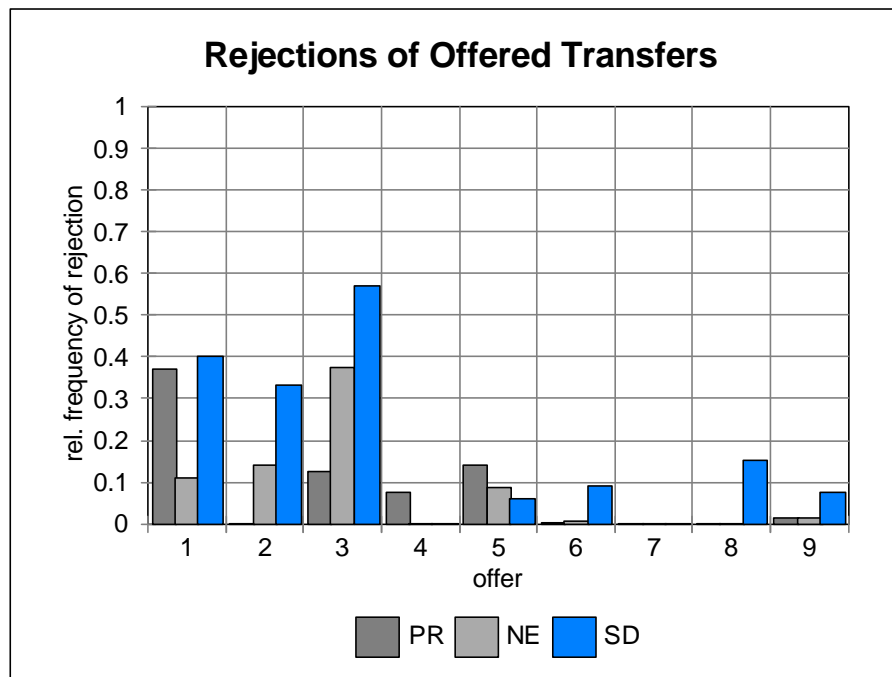


Figure 6

5.4. The Probability Estimations

Finally, we test whether hypothesis 4, which claimed that subjects tend to under-estimate the probability of being penalised, can be supported by the results of the final questionnaires on probability estimations. The answers of the questionnaire expose a very clear picture. Figure 7 shows the average estimates for the nine different tasks. It can be seen that, first, the average estimates are typically well below the correct answers, and, second, subjects respond to the difference of the probabilities, but not very strongly to the difference in the round number.

24 of the 27 questionnaires that were examined exhibit more under- than over-estimations¹⁵, two show an equal number of both, and only one subject over-estimated the overall probability of disqualification more often. In all eleven pairs for which usable estimates are available for both players, the sum of all estimates of both players was lower than the sum of the answers that would have been correct. The binomial test rejects the null-hypothesis of equal probability of positive and negative differences between the sum of estimates and the sum of the correct answers in favour of the alternative hypothesis of more under-estimations at a significance level of $\alpha = 0.001$ (one-tailed).

It is interesting that although subjects tend to under-estimate the probability of the sudden death over the whole experiment, the penalty still shows a strong effect towards lower transfers and

¹⁵ Since one pair was disqualified, only 34 of the 36 participants filled in the questionnaire. For security, we excluded 7 sets of answers by subjects who were possibly confused about single round and overall probabilities, as they typed in the former as answers. Note that this outlier rejection is conservative to our findings.

less other-damaging choices. This strengthens the result that the fear of a drastic penalty has a strong effect in reducing corrupt behaviour.

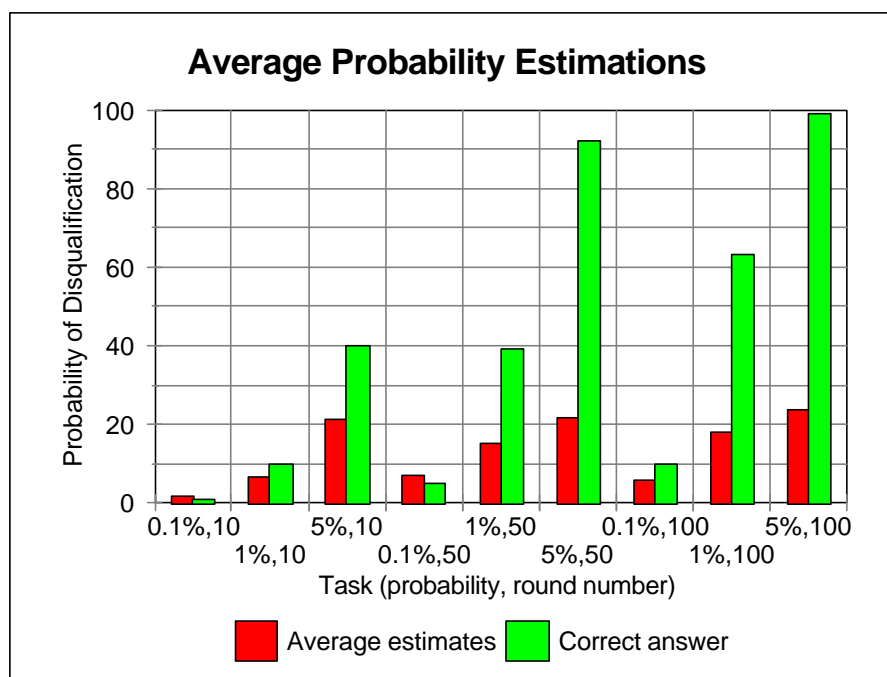


Figure 7

A correlation between the individual tendency to under-estimate the disqualification probability, and the tendency to pay or accept bribes cannot be detected. However, since there are two players needed to establish a bribery relationship, possible effects in that sense might not show up in the data, since two players with different estimation patterns might be matched together.

Observation 4. *Subjects significantly tend to underestimate the overall probability of disqualification.*

6. Summary and Conclusions

We introduced an experimental bribery game to separate three major characteristics of a bribery relationship: the reciprocity character, the negative external effect to others, and the threat of severe penalties. We analysed corruption as a long-term relationship in which the same briber meets the same public official again and again. Three treatments were conducted. In the basic treatment, a simple reciprocity game was played as a control experiment. We find that trust and reciprocation can establish stable exchange of benefits, although this is off the equilibrium. Reciprocal behaviour is exhibited by both players: public officials reciprocate on sufficiently high transfers, but also firms condition their transfers on the observed decisions by the officials in the previous round. In a second treatment, we introduced a negative externality that was imposed to all other subjects in case that the reciprocal response was chosen. The results

were strikingly similar to the basic treatment. Both the distributions of transfers and their evolution over time are not apparently different from those in the pure reciprocity treatment. The occurrence of corrupt behaviour is rising in the beginning. Corruption reaches a stable high level after a co-ordination period.

To examine the effect of severe penalties that are imposed with very low probability, we conducted a third treatment in which we introduce a new feature which we call the sudden death treatment. Whenever a bribe is accepted, a lottery is played out, and in the negative outcome both players are excluded from the experiment without payment. We find that, although the discovery probability was very low and the expected value of reciprocation still greater than that of equilibrium play, the threat of the sudden death penalty induced significantly lower transfers and frequencies of other-damaging choices. Further, the number of rejected offers is increased. It is interesting that the sudden death penalty takes a significant effect although our questionnaires show that the overall probability of getting disqualified in the session is typically underestimated. The latter result supports a hypothesis posed by KIRCHGÄSSNER (1997).

Though our study was designed to fundamentally examine behaviour in a corruption situation, the results might be of practical relevance for the combat against bribery. When designing public relations campaigns against bribery, it is crucially important to know which are the most important behavioural influences. Our results cast doubt on the effectiveness of campaigns appealing to the consciousness of officials of the negative welfare effects of bribery. On the other hand, our data suggest to emphasise the threat of getting caught, especially they suggest to aim at correcting the widespread miscalibration of discovery probabilities. Given a strong negative effect of drastic penalties on corruption levels, we can expect that individuals would then be even more reluctant to take the risk of corruption.

References

- ABBINK, Klaus, Bernd IRLBUSCH, and Elke RENNER (forthcoming): The Moonlighting Game – An Experimental Study on Reciprocity and Retribution. *Journal of Economic Behavior and Organization*.
- ABBINK, Klaus, and Abdolkarim SADRIEH (1995): RatImage - Research Assistance Toolbox for Computer-Aided Human Behavior Experiments. SFB Discussion Paper B-325, University of Bonn.
- BERG, Joyce, John DICKHAUT, and Kevin McCABE (1995): Trust, Reciprocity and Social History. *Games and Economic Behavior*, 10, 122-142.
- BURNHAM, Terrence, Kevin McCABE, and Vernon L. SMITH (1998): Friend- or Foe-Intentionality Priming in an Extensive Form Trust Game. Discussion Paper, University of Arizona.
- DAVIS, Suzanne (1997): The Art of Corporate Gift Giving. *Incentive*, 3-40.
- DUFWENBERG, Martin, and Uri GNEEZY (forthcoming): Efficiency, Reciprocity, and Expectations in an Experimental Game. *Games and Economic Behavior*.
- FAHR, René, and Bernd IRLBUSCH (forthcoming): Fairness as a Constraint on Trust and Reciprocity – Earned Property Rights in a Reciprocal Exchange Experiment. *Economics Letters*.
- FEHR, Ernst, Simon GÄCHTER, and Georg KIRCHSTEIGER (1997): Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica*, 65:4, 833-860.
- FEHR, Ernst, Georg KIRCHSTEIGER, and Arno RIEDL (1993): Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics*, 108, 437-459.
- FRANK, Björn, and Günther G. SCHULZE (1998): How Tempting is Corruption – More Bad News About Economists. Working Paper, Universität Hohenheim.

- GUPTA, Sanjeev, Hamid DAVOODI, and Rosa ALONSO-TERME (1998): Does Corruption Affect Inequality and Poverty? IMF Working Paper WP/98/76.
- HOFFMAN, Elizabeth, Kevin A. MCCABE, and Vernon L. SMITH (1998): Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology. *Economic Inquiry*, XXXVI, 335-352.
- JACOBSEN, Eva, and Abdolkarim SADRIEH (1996): Experimental Proof for the Motivational Importance of Reciprocity. SFB Discussion Paper B-386, University of Bonn.
- KAHNEMAN, Daniel, Jack L. KNETSCH, and Richard H. THALER (1986a): Fairness and the Assumptions of Economics. *Journal of Business*, 59:4, S285-S300.
- KAHNEMAN, Daniel, Jack L. KNETSCH, and Richard H. THALER (1986b): Fairness as a Constraint on Profit-Seeking. *American Economic Review*, 76:4, 728-741.
- KIRCHGÄSSNER, Gebhard (1997): Auf der Suche nach dem Gespenst des Ökonomismus – Einige Bemerkungen über Tausch, Märkte und die Ökonomisierung der Lebensverhältnisse. *Analyse und Kritik*, 19, 127-152.
- KLITGAARD, Robert (1988): Controlling Corruption. University of California Press.
- LEDYARD, John O. (1995): Public Goods. In John Kagel and Alvin E. Roth (eds): The Handbook of Experimental Economics. Princeton University Press.
- LONG, D. Michael, and Spuma RAO (1995): The Wealth Effects of Unethical Business Behavior. *Journal of Economics and Finance*, 19:2, 65-73.
- LÜDTKE, Hartmut, and Hartmut SCHWEITZER (1993): Korruptionsneigung bei unterschiedlichen Erwartungskonstellationen in der Handlungssituation – Ein Quasi-Experiment mit Studenten. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 45:3, 465-483.
- MANION, Melanie (1996): Corruption by Design: Bribery in Chinese Enterprise Licensing. *The Journal of Law, Economics, & Organization*, V12N1, 167-195.
- MAURO, Paolo (1995): Corruption and Growth. *Quarterly Journal of Economics*, 680-712.
- MEIDINGER, Claude, Stéphane ROBIN, and Bernard RUFFIEUX (1998): A repeated investment game. Mimeo, Ecole Nationale Supérieure du Génie Industriel, Grenoble.
- NECKEL, Sighart (1995): Der unmoralische Tausch – Eine Soziologie der Käuflichkeit. In: Karl Markus Michel and Tilman Spengler (eds.): Korruption. Rowohlt Verlag, Berlin.
- OFFER, Avner (1997): Between the Gift and the Market: The Economy of Regard. *Economic History Review*, 50:3, 450-476.
- PRUITT, D.G. (1970): Reward Structure of Cooperation: The Decomposed Prisoners' Dilemma Game. *Journal of Personality and Social Psychology*, 7, 21-27.
- ROSE-ACKERMAN, Susan (1975): The economics of corruption. *Journal of Public Economics*, 4, 187-203.
- ROSE-ACKERMAN, Susan (1985): Corruption – A Study in Political Economy. Academic Press.
- SELTEN, Reinhard (1978): The Equity Principle in Economic Behavior, in: Decision Theory and Social Ethics - Issues in Social Choice, D. Reidel Publishing Company, Dordrecht, pp. 289-301.
- SELTEN, Reinhard, and Rolf STOECKER (1986): End Behavior in Sequences of Finite Prisoners' Dilemma Super-games – A Learning Theory Approach. *Journal of Economic Behavior*, 47-70.
- SHLEIFER, Andrei, and Robert W. VISHNY (1993): Corruption. *Quarterly Journal of Economics*, 108, 599-617.
- TANZI, Vito (1998): Corruption Around the World: Causes, Consequences, Scope, and Cures. IMF Working Paper WP/98/63.
- VAN RUCKEGHEM, Caroline, and Beatrice WEDER (1997): Corruption and the Rate of Temptation: Do Low Wages in the Civil Service Cause Corruption? IMF Working Paper WP/97/73.

Appendix A. Proof of the Proposition.

To keep the proof general for various payoff parameter constellations, define the following variables (the values in brackets are the parameters used in the experiment). Following the sequence of moves in the game, we call the firm player 1, and the public official player 2.

B^{1X}	player 1's payoff if X is chosen	[36]
B^{1Y}	player 1's payoff if Y is chosen	[56]
B^{2X}	player 2's payoff if X is chosen	[36]
B^{2Y}	player 2's payoff if Y is chosen	[30]
f	transfer fee	[2]
t^{max}	maximum transfer	[9]
2	sudden death probability per round	[0.003]

where the inequalities $B^{1X} < B^{1Y}$, $B^{2X} > B^{2Y}$, $0 \neq 2 \neq 1$, $f > 0$, and $t^{max} > 0$ hold. Further, denote by $t(I^1)$ player 1's transfer at an information set I^1 , by $d(I^2)$ player 2's choice of X or Y at information set I^2 , and by $a(I^2)$ player 2's decision at an information set I^2 at which she has to decide on accepting or rejecting. For convenience, we use a subscript to denote the round at which this information set is reached, such that I_j^1 is an information set of player 1 for round j , and by I_j^2 an information set at which player 2 decides on X or Y in round j . An information set contains the history of play of the two players directly matched to one another.

Since the payoffs the players gain in the game are additively composed of payoffs resulting from their own moves and the damages done to them by others, we may define $A^i(I_j^k)$ as player i 's cumulative payoff earned through the moves of the players of the own pair, up to the point of time when I_j^k is reached, and by $*(s^{1,1,2})$ the (expected) damages done to each player of a pair by the other players' (mixed) strategies $s^{1,1,2}$. For convenience, we drop the argument if it is not required.

Denote by $p(I_j^i)$ the probability with which information set I_j^i is reached. Denote by q the probability with which a player makes a certain decision at a certain information set. Let us first show that in equilibrium Y is never chosen in the last round. Consider some information set I_n^j , where n is the last round. Suppose there is an equilibrium $E = (s^1, \dots, s^n)$ with $q(d(I_n^2) = Y) > 0$ and $p(I_n^2) > 0$. Now consider player 2's alternative strategy which differs from his strategy in E only in $d(I_n^2) = X$. $p(I_n^2)$ remains unchanged. Player 2's payoff with the alternative strategy is $A^2(I_n^2) + B^{2X} + *$ which is higher than player 2's payoff in E , that is $A^2(I_n^2) + B^{2Y}q(d(I_n^2) = Y) + B^{2X}(1 - q(d(I_n^2) = Y)) + *$. Thus, E cannot be an equilibrium.

Now let us show that in equilibrium, player 1 will never transfer a positive amount in the last round. Suppose there is an equilibrium E with $q(t(I_n^1) = 0) < 1$ and $p(I_n^1) > 0$. Since E is an equilibrium, player 2 will not choose Y in the following stage of the last round. Thus, the best possible payoff player 1 can get in an equilibrium characterised by $q(t(I_n^1) = 0) < 1$ and $p(I_n^1) > 0$ is $A^1(I_n^1) + B^{1X}q(0) + (B^{1X} - f)(1 - q(0)) + *$. This is the case if in E , player 2's strategy involves rejecting the transfer offers made by player 1. If player 2's strategy in E would involve accepting transfer offers made by player 1, player 1's payoff can only be lower, since it is further reduced by transfers and the possibility of sudden death.

Whatever equilibrium strategy player 2 would choose, it can easily be seen that player 1's payoff by choosing the alternative strategy, which differs from his strategy in E only in that $q(t(I_n^1) = 0) = 1$, is always higher than his payoff in E . All other things are equal, $p(I_n^1)$ remains unchanged. Player 1's payoff in case that the information set is reached, however, is in the range from $A^1(I_n^1) + B^{1X} + *$ (player 2 chooses X with probability 1 on all paths reached through player 1's alternative strategy and player 2's strategy in E) to $A^1(I_n^1) + B^{1Y} + *$ (player 2 chooses Y for sure). Since all possible payoffs are higher than the maximum payoff in an equilibrium E , given in E $q(t(I_n^1) = 0) < 1$ holds, E cannot be an equilibrium.

The proposition can now be proved by mathematical induction. Denote by $S = \{j, \dots, n\}$ a set of consecutive rounds for which in equilibrium $d(I_j^2) = \dots = d(I_n^2) = X \text{ \ae } I_i^2$ with $p(I_i^2) > 0$, $i = j, \dots, n$, and $t(I_j^1) = \dots = t(I_n^1) = 0 \text{ \ae } I_i^1$ with $p(I_i^1) > 0$, $i = j, \dots, n$. We have shown that $n \notin S$. Now consider round $j + 1$. Suppose there is an equilibrium E in which for some information set $p(I_{j+1}^2) > 0$ and $q(d(I_{j+1}^1) = Y) > 0$. Since E is an equilibrium, player 2's payoff in case that I_{j+1}^2 is reached is $A^2(I_{j+1}^2) + B^{2Y}q(d(I_{j+1}^2) = Y) + B^{2X}(1 - q(d(I_{j+1}^2) = Y)) + B^{2X*}S^* + *$. Note that since E is an equilibrium, there will be no positive transfers in later rounds. Further, since E is an equilibrium, player 2's strategy must be a best reply to all other players' strategies. It is now sufficient to show that player 2 has at least one alternative strategy which is a better reply to all other players' strategies in E . Consider, for example, player 2's alternative strategy that differs from her strategy in E in that $d(I_i^2) = X$ and $a(I_i^2) = \text{reject} \text{ \ae } i \notin S$. All other things equal, $p(I_j^2)$ remains unchanged, but player 2's payoff given I_j^2 is reached is $A^2(I_j^2) + B^{2X}(*S^* + 1) + *$, which is higher than in E . Thus, E cannot be an equilibrium.

Now suppose there is an equilibrium E in which $q(t(I_j^1) = 0) < 1$ and $p(I_j^1) > 0$ for some information set I_j^1 . Since E is an equilibrium, $d(I_k^2) = X \text{ \ae } k \notin j + 1$ and $p(I_k^2) > 0$, and $t(I_k^1) = 0 \text{ \ae } k \notin j$ and $p(I_k^1) > 0$. Thus player 1's expected payoff, given I_j^1 is reached, and given $q(t(I_{j+1}^1) = 0)$, is in the range from $A^1(I_j^1) + B^{1X}(*S^* + 1)q(0) + (1 - 2)(1 - q(0)) \text{ \ae } (B^{1X} - f - t^{max} + B^{1X}(*S^* + 1)) + *$, if player 1 transfers the maximum with probability 1, and player 2 accepts for sure, to $A^1(I_j^1) + B^{1X}(*S^* + 1)q(0) + (1 - q(0))(B^{1X} - f + B^{1X*}S^*) + *$, if player 2 rejects all positive amounts transferred. However, since E is an equilibrium, player 1's strategy must be a best reply to all other players' strategies. It is now sufficient to show that player 1 has at least one alternative strategy which is a better reply to all other players' strategies in E . Consider, for example, player 1's alternative strategy that differs from the equilibrium strategy in that $q(t(I_k^1) = 0) = 1 \text{ \ae } I_k^1$ with $k \notin j$. Player 1's payoff, given I_j^1 is reached, is then in the range from $A^1(I_j^1) + B^{1X}(*S^* + 1) + *$, if player 2 chooses X in all later rounds with probability 1, to $A^1(I_j^1) + B^{1Y}(*S^* + 1) + *$, if player 2 always chooses Y for certain. No matter what player 2's strategy is off the equilibrium path of E , the alternative strategy is a better response to all other players' strategies in E than player 1's strategy in E . Thus, E cannot be an equilibrium. This implies that $j \notin S$. \text{ \ae }

The proof holds for all three treatments, since the negative externality and the pure reciprocity treatments are special cases with $2 = 0$, $* = 0$ and $2 = 0$ respectively.

Appendix B. The Instructions for the Sudden Death Treatment

(original text in German, other treatments analogous)

All in all **18 persons** participate in the decision making experiment. There are two types of participants: **Player 1** and **Player 2**. At the beginning of the experiment, the type of each participant is randomly drawn. **The type of a participant remains unchanged throughout the experiment.**

Pairs of players are matched also randomly: One player 1 and one player 2 are matched to one another. **The pairs remain unchanged throughout the experiment.**

The experiment consists of **30 rounds**. At the end of the experiment you will receive a **payoff** that depends on your success.

Decision Situation in a Round

Stage 1: Transfer or no Transfer

First, player 1 decides whether or not he wants to transfer an amount to player 2. If he does, then the credit of player 1 is reduced by **2 talers**, and the play is continued with stage 2. If player 1 does not want to transfer an amount, then both credits remain unchanged, and the play is continued with stage 4.

Stage 2: The Amount to Be Transferred

Player 1 decides on the amount to be transferred to player 2. Player 1 can choose between **1, 2, 3, 4, 5, 6, 7, 8 or 9 talers**. The play is continued with stage 3.

Stage 3: Acceptance or Rejection of the Transfer

Player 2 decides on whether he **accepts** or **rejects** the proposed transfer.

- If player 2 accepts the transfer, then the credit of player 1 is reduced by the amount he proposed. Player 2's credit is increased by the **tripled amount** that is transferred. In the following, a number out of the range from 0 to 999 is randomly drawn.
 - ⇒ If the randomly drawn number is **0, 1, or 2**, then player 2 and the player 1 matched with him are **disqualified**. That means: **The play ends for these two players, and they do not receive any payment for the play, i.e. also the talers that have been earned in the past are cleared from their accounts.** (In the end of the experiment, both players receive only the show up fee, see below). The two disqualified participants fill in a questionnaire, until the experiment has ended. For the other participants, the play is continued normally.
 - ⇒ If the randomly drawn number is **3, 4, ..., 998, or 999**, then the play is continued with stage 4.
- If player 2 rejects the transfer, then the credits remain unchanged (The transfer fee from stage 1, however, is also paid in case of rejection). The play is continued with stage 4.

Stage 4: Choice Between X and Y

Player 2 chooses one of the **alternatives X or Y**.

- If player 2 selects alternative X, then his credit and the credit of the player 1 matched with him are increased by **36 talers** each. The credits of the 16 other participants is **not changed** by this decision.
- If player 2 selects alternative Y, then player 1's credit is increased by **56 talers**, whereas player 2's credit is increased by **30 talers**. The credit of each of the 16 other participants is **decreased by 3 talers** by this decision.

Attention: by each of the eight other pairs, in which Y is chosen, the payoff for player 1 as well as for player 2 is decreased by 3 talers, i.e. at maximum eight times 3 and at minimum no talers are deducted from player 1's and player 2's credits each. The deductions by decisions of other pairs are not announced before the experiment has ended.

After stage 4, the round has ended. The round payoffs are the sum of all changes of credits during the four stages of the round.

The payoffs

You receive your payoff at the end of the experiment, where the exchange rate is DM 3.00 for 100 talers. In addition, you receive a lump sum show up fee of DM 5.00.

The Payoff Tables

Round payoff if player 2 **accepts** a transfer

Transferred amount	1		2		3		4		5		6		7		8		9	
Player 2's decision	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
Payoff ... Player 1	33	53	32	52	31	51	30	50	29	49	28	48	27	47	26	46	25	45
... Player 2	39	33	42	36	45	39	48	42	51	45	54	48	57	51	60	54	63	57
... each of the other 16 participants	0	-3	0	-3	0	-3	0	-3	0	-3	0	-3	0	-3	0	-3	0	-3

Round payoff if player 2 **rejects** a transfer

Transferred amount	1,...,9	
Player 2's decision	X	Y
Payoff... ... Player 1	34	54
... Player 2	36	30
... each of the 16 other participants.	0	-3

Round payoff if player 2 does **not** transfer an amount

Transferred amount	0	
Player 2's decision	X	Y
Payoff... ... Player 1	36	56
... Player 2	36	30
... each of the 16 other participants.	0	-3

Each of the 16 other pairs in which Y is chosen decreases the payoff for player 1 and player 2 by another 3 talers each.

The Main Screen Display of the Sudden Death Treatment

