

Schikora, Jan

**Conference Paper**

## Bringing the Four-Eyes-Principle to the Lab

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Experimental Tests of Theories and Institutions, No. D13-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Schikora, Jan (2010) : Bringing the Four-Eyes-Principle to the Lab, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Experimental Tests of Theories and Institutions, No. D13-V2, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37465>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Bringing the Four-Eyes-Principle to the Lab

Jan Theodor Schikora\*

August 2, 2010

## **Abstract**

The ‘Four-Eyes-Principle’ is considered as one of the most potent measures against corruption although it lacks both theoretical and empirical justification. We show in a laboratory experiment that the effect of introducing the 4EP is negative. This result is contrary to theoretical predictions derived in a standard model of corruption. Combining data of choices with results of a content analysis of exchanged messages we show that the individual profit maximizing motive dominates the group decision making process despite the trade-off with social efficiency. Against existing policy recommendations, the results of our experiment cast doubt on the usefulness of the 4EP.

**Keywords:** Corruption; Laboratory Experiments; Group Decision Making

**JEL Classification Numbers:** C72. C92. D73

---

\*MGSE, University of Munich

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Experimental set-up</b>	<b>6</b>
2.1	Corruption and the NIE . . . . .	6
2.2	Introducing the 4EP . . . . .	8
<b>3</b>	<b>Hypotheses</b>	<b>11</b>
3.1	Bribe Splitting Effect . . . . .	12
3.2	Group decision-making Effect . . . . .	14
3.2.1	Competitive Motive . . . . .	16
3.2.2	Profit Maximizing Motive . . . . .	16
3.3	<i>B</i> 's behaviour . . . . .	17
3.4	Gender Effects . . . . .	18
3.5	Total effect . . . . .	19
<b>4</b>	<b>Procedure</b>	<b>20</b>
<b>5</b>	<b>Results and Interpretation</b>	<b>21</b>
5.1	Descriptive results . . . . .	22
5.1.1	Corruption levels . . . . .	23
5.1.2	Payoff and externalities . . . . .	24
5.1.3	Transfer levels . . . . .	27
5.2	Conditional reciprocity . . . . .	29
5.2.1	Conditional reciprocity of <i>O</i> . . . . .	29
5.2.2	Conditional reciprocity of <i>B</i> . . . . .	33
5.3	Switching behaviour and content analysis . . . . .	34
5.3.1	Switching behaviour . . . . .	35
5.3.2	Content Analysis . . . . .	36
<b>6</b>	<b>Conclusion</b>	<b>38</b>

## List of Figures

1	Extensive form representation of the 3-Stage Game . . . . .	8
2	Monetary Costs and Benefits of corruption in TDT1 and TDT2 . . . . .	13
3	Success-probabilities over Periods . . . . .	24
4	Payoff levels and levels of externalities . . . . .	25
5	Relative distribution of transfer levels . . . . .	28
6	Success rates conditional on Transfer . . . . .	30
7	Extensive forms of TDT1 and IDT1 . . . . .	42
8	Extensive forms TDT2 and IDT2 . . . . .	42

## List of Tables

1	<b>Performance variables</b> . . . . .	23
2	<b>Gini Coefficients</b> . . . . .	25
3	<b>Output of the Linear Probability Model M3</b> . . . . .	31
4	<b><i>R</i>-value and adjusted <i>R</i>-value</b> . . . . .	34
5	<b>Success and Initial Consent, TDT1</b> . . . . .	35
6	<b>Success and Initial Consent, TDT2</b> . . . . .	35
7	<b>Success and Content</b> . . . . .	37
8	<b>Model M1</b> . . . . .	43
9	<b>Model M2</b> . . . . .	43
10	<b>Model M4</b> . . . . .	44

# 1 Introduction

With almost daily media attention of high profile scandals, corruption, generally defined as ‘the misuse of public office for private gain’ (OECD 2010), has been recognized as a major problem. In general, a corrupt transaction is illegal, since it exerts a large negative external effect on outsiders, which is usually larger than the sum of benefits to the agents involved. This defines corruption as socially inefficient.

In addition to traditional views of deterring an agent from engaging in criminal activity by varying the amount of penalties and the probability of detection (Lambsdorff 2007), the New Institutional Economics (NIE) of corruption concentrates on finding a fine-tuned set of policy tools defining an institutional design that optimally exploits the instability of the corrupt deal between a client and an official, (Lambsdorff and Nell 2005).<sup>1</sup>

The instability of a single corrupt transaction stems from the enforcement problem between a bribing agent and a potentially corrupt official (Klitgaard 1988, Lambsdorff 2005, 2007) which can be linked to its illegal nature precluding the assistance of legal third parties. The occurrence of corruption relies heavily on trust and reciprocity between the two agents and is difficult to explain in standard economic models. Nonetheless, national as well as international organizations such as Transparency International, the OECD and several national fiscal authorities publish lists of (institutional) policy recommendations containing measures to curb corruption.

Along with ‘staff rotation’ (studied in Abbink 2004), the introduction of the 4EP, ‘a requirement that business has to be effectively conducted by at least two individuals (four eyes)’ is one of the most prominent examples (Poerting and Vahlenkamp 1998, Rieger 2005, Wiehen 2005).

As a result of a general lack in the theoretical tractability (let alone predictability) of actual behaviour in the corrupt context<sup>2</sup>, there is no fully-fledged theoretical analysis of its effectiveness. Nor is there any kind of traceable empirical evidence to support the usefulness of the introduction of the 4EP.

---

<sup>1</sup>There is a considerable amount of theoretical research on the Principal-Agent relationship between a (benevolent and non-corrupt) government and its public officials (Becker 1968, Groenendijk 2004). Ignoring the lack of (legal) enforcement of a corrupt transaction between  $O$  and  $B$  boils the problem down to the analysis of an ordinary Principal Agent model with an interesting application.

<sup>2</sup>In contrast to theoretical predictions applying the standard self interested model within the standard game of corruption (see Section 2), Abbink et al (2002), Abbink (2004) and Lambsdorff and Frank (2007) find a large amount of cooperation between corrupt partners in a series of laboratory experiments using one-shot as well as (finitely) repeated versions of the standard game. Considering, that the negative externalities make any successful corrupt deal socially inefficient, these levels of reciprocity cannot be easily explained by models of ‘strong reciprocity’ (Fehr and Schmidt 2006) nor any other model allowing for any meaningful predictions.

Even on a more general level, the distinction between individuals and small groups as decision-makers has been widely ignored in the theoretical literature. The problem of differences between the behaviour of individuals deciding alone or in a group has only recently been addressed in the field of experimental economics, where results seem ambiguous. Some studies find that groups tend to behave closer to standard equilibrium predictions derived from the self interested model of payoff maximization (e.g. Bornstein and Yaniv 1998, Blinder and Morgan 2005), other studies (e.g. Kocher and Sutter 2005, 2007, Cason and Mui 1997) provide experimental evidence of the contrary. Kocher and Sutter (2007) conclude from their studies that the direction of the group decision making effect critically depends on the nature of the task determining which of two countervailing motives, the profit maximizing motive or the competitive motive, dominates.

The basis of our experimental set-up is close to the ones used in Abbink et al (2002, 2004) and Lambsdorff and Frank (2010). Our experiment is designed to assess the effects of the introduction of the 4EP on observed levels of corruption. Within this framework we model the 4EP as replacing a single official (deciding individually) with a group of two officials deciding jointly according to a decision making process that secures veto power for the non-corrupt official. Despite considering the entity of only two participants we call it a group rather than a team.

With four different treatments we can separate two countervailing effects of the introduction of the 4EP. One is due to the change in marginal incentives resulting from the division of the transfer between the members of the decision making unit. The other effect is determined by the group decision making process alone (keeping marginal incentives constant). Rejecting predictions taken from (self interest based) arguments within the standard model of corruption, we find that the introduction of the 4EP increases the frequency of successful corrupt transactions unambiguously. We argue that this is due to the dominance of the profit maximizing motive (Kocher and Sutter 2007). By their higher (joint) cognitive capacity, groups of officials seem to be more capable of maximizing their profits by following strategies that are shown to lead to a higher frequency of corrupt transactions based on mutual reciprocity.

The differences in revealed strategies, i.e. more accentuated behaviour with respect to conditional responding, by groups (compared to individuals) can be identified by the analysis of conditional corruptibility. The results of a (rudimentary) content analysis of chat messages exchanged during the decision making process within groups are in line with the explanations

of the observed effects derived on the basis of the persuasive argument theory (Pruitt 1971). Moreover, the content analysis does not only substantiate the dominance of the ‘profit maximizing motive’ (Blinder and Morgan 2005) it also provides an important insight into the black box of the mechanisms of group decision making. Since groups perform better in solving the enforcement problem between briber and official, the introduction of the 4EP moves behaviour further away from the theoretical predictions of the social optimum. Therefore our results cast doubt on the usefulness of the introduction of the 4EP and its justification as a recommended measure against corruption.

The remainder of the paper is structured as follows. Section 2 describes the experimental set-up giving details on the specifications of all four treatments. Section 3 analyses the effects of the introduction of the 4EP in the framework of the NIE of corruption and forms hypotheses based on simple arguments within expected utility theory. Section 4 describes the main findings and gives a detailed explanation of the empirical strategies and interprets the results. Section 5 concludes.

## 2 Experimental set-up

### 2.1 Corruption and the NIE

In its most general form, a corrupt transaction can be described as a Principal-Agent-Client relationship, in which the principal, represented by the government or any kind of benevolent<sup>3</sup> authority, deals with its clients (private entities, e.g. individuals or firms) through its agents, the potentially corrupt (not perfectly monitored) public officials. Clients may have an incentive to transfer side-payments (i.e. bribes) to relevant officials in order to alter their behaviour with respect to their duties (i.e. fulfilling legal procedures clearly defined but not perfectly controlled by the authorities).

The main mechanism of the instable and therefore interesting part of the relationship is best explained in a simple 3-Stage game (Lambsdorff 2005, 2007):

In Stage 1, a client (e.g. a potentially bribing firm)  $B$  decides on the level of bribe  $b=[0;12]$  Experimental Monetary Units (EMU) to be given to a potentially corrupt official  $O$ . The limit of 12 EMU reflects  $B$ ’s budget constraint per period. At this point the amount of  $b$  is tripled. The factor 3 captures the idea of a large difference in the marginal utilities of money between a (wealthy) client and a (poor) official (Abbink et al 2002).

---

<sup>3</sup>This excludes some cases of political or ‘grand’ corruption (see Klitgaard 1988).

In Stage 2.  $O$  can either ‘accept’ the transfer  $b$  and keep the trippled amount ( $3*b$ ) to himself or ‘reject’ it. In the latter case the game ends,  $O$  keeps only  $b$  to himself (as some ‘benefit’ from pro social behaviour) while the rest (the remaining amount of  $2*b$  is used for the ‘public benefit’.<sup>4</sup> In our experiment, the ‘public’ is modelled in two different ways. In one set of sessions (mode 1) we model the externality on the public as payments (reductions of payments) to four randomly chosen participants of the (relevant session of the) experiment. For technical reasons a particular subject can never be hit more than once by an external effect in one period.

In the second set of sessions (mode 2) we model the externalities as increases or decreases in the amount of a donation to the public aid organization ‘Doctors without Borders’.<sup>5</sup> The total amount of added or deducted payments is equal across the two modes. Using two different modes of modelling the externality including real outsiders allows us to address the problem of a ‘super-game’ considering the possibility of participants forming expectations on the behaviour of participants outside their own group.

If  $O$  accepts the bribe, she gets the trippled transfer for sure. She automatically enters Stage 3 where she decides between two options. The first option includes initiating an increase of  $B$ ’s payoff by 16 EMU (delivering the corrupt service) at the fixed costs of 4 EMU (to himself). In this case, the ‘public’ suffers substantially (by  $-24$  EMU). Note that independent from the size of the transfer, the negative externality is always larger than the sum of payoffs for the agents  $B$  and  $O$  so that a successful corrupt transaction is always inefficient.<sup>6</sup> In the second option  $O$  arranges nothing (implicitly defects on  $B$ ), saving costs for himself and the public from the negative externality. The costs of delivery may be interpreted either as the sum of the certainty equivalent of the lottery between punishment in the case of detection and undetected corrupt success, or as the practical costs of engaging in a criminal activity, e.g. hiding illegal activities from colleagues and superiors.<sup>7</sup>

If  $O$  decides against delivering the corrupt service in Stage 3.  $B$  does not get the bonus,  $O$  does not bear the costs of 4 EMU and only a minor negative externality (4 EMU in total) is

---

<sup>4</sup>This implies the assumption of equal marginal utility between  $O$  and the public.

<sup>5</sup>We chose ‘Doctors without Borders’ to produce results as comparable as possible to the findings of Lambsdorff and Frank (2007) who use this organization in their experiment. As expected we find from answers to our post-experimental questionnaire that virtually all subjects approve of this organization and take this as evidence for our working hypothesis that the reduction of a real donation represents a valid model for the reduction of public welfare.

<sup>6</sup>See Bardhan (1997) and Klitgaard (1988) for a review of (empirical) evidence supporting the assumption of the negative externality and inefficiency of corruption.

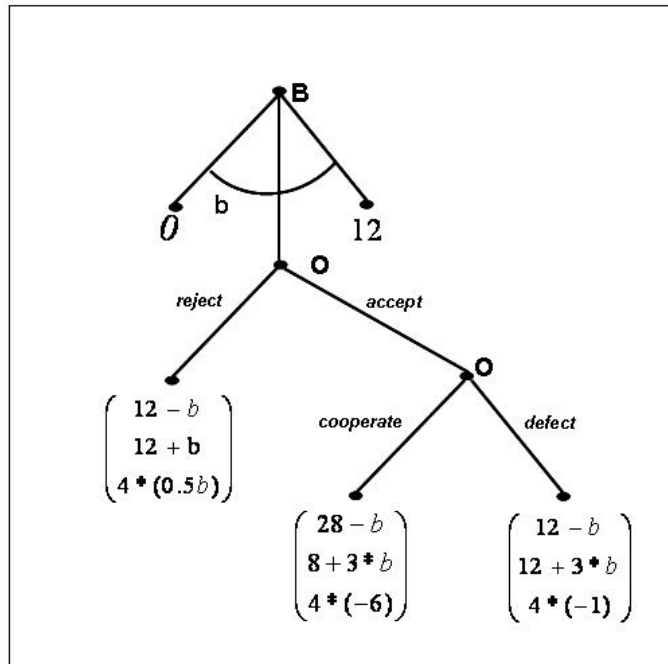
<sup>7</sup>Some theoretical and experimental studies (e.g. Abbink 2004) use a fourth stage, in which there is a probability of detection. Using a lottery instead of a fixed amount to model the cost of corruption is problematic since this adds another problem of individual differences in risk aversion. It would be necessary to disentangle potential treatment effects from differences in risk aversion at this level.



realized. Within the standard self interested model it is never optimal, neither in a one shot nor in a finitely repeated game, to deliver at a personal cost, see Appendix A for a short proof.

Figure 1 represents the time structure of the basic set-up in its extensive form.

Figure 1: Extensive form representation of the 3-Stage Game



## 2.2 Introducing the 4EP

In order to investigate the effects of the introduction of the 4EP into this set-up, we consider two participants instead of one making the decisions of the official in Stage 2 and Stage 3.

In order to assess a possible effect from the introduction of the 4EP on the level of corruption, we define the level of corruption as the frequency of successful corrupt deals relative to the total number of possible transactions.<sup>8</sup>

We use four different versions of the game as treatments of the experiment. In all four treatments, subjects play their version of the corruption game for ten successive periods. Af-

<sup>8</sup>Note that in our set-up this number is directly proportional to the sum of negative externalities, but not directly proportional to the average (or total) payoff of the participants.

ter each period they learn about their group partners’/partner’s (payoff relevant) choices.<sup>9</sup> All treatments are run in a partner design so that all subjects remain in their respective unit (of  $B$  and  $O$  subjects) for all the ten periods of the experiment. Separating potentially countervailing effects and still being able to compare outcomes across treatments, we have to consider different group sizes within the treatments. Find Figures 7 and Figure 8 in Appendix B for the full extensive forms of the games played in the respective treatments.

### **IDT1**<sup>10</sup>

In Treatment 1, the ‘Individual Decision-making Treatment 1’ (IDT1), we consider units of two subjects, one in the role of the official  $O$  and one in the role of the potential briber  $B$ . The 3-Stage corruption game (see Figure 1) is played for ten consecutive periods. At the end of each period, all participants get to know their own payoffs. Additionally, type  $B$  subjects get information about the (Stage 2 and Stage 3) decisions of their transaction partners (of type  $O$ ). While through this information all subjects know about the negative or positive externalities they have contributed to cause to the public (four randomly chosen participants in mode 1 or ‘Doctors without Borders’ in mode 2) they do not learn about the magnitude of the spillovers that may have been caused to them by the decisions of the subjects outside their unit in mode 1. Nevertheless, there could be violations of the independence assumption due to considerations of a ‘super-game’. Participants may condition their choices on their beliefs about the other participants’ corrupt behaviour which may affect them through the negative externality. To check for the existence of effects stemming from this ‘super-game’ and provide a robustness check, we use mode 2 in which the public is modelled as a ‘real’ outsider, the recipients of the donation towards the public aid organisation. At least for this mode independence of observations across units is warranted.

### **TDT1**

In the second treatment, the ‘Team Decision-making Treatment 1’ (TDT1), we form units of three subjects, one  $B$  and two  $O$  types. The  $B$  type decides in Stage 1 about her bribe  $b$  which is trippled and then transferred to both officials of her (3-player-)unit. Note that although the amount goes to two players, it is subtracted only once from  $B$ ’s account. The parameters of the game are set in such a way that the incentives for the officials are

---

<sup>9</sup>By choosing a repeated instead of a one-shot set-up, we focus on the strategic component of the reciprocal transaction of corruption which we expect to be affected by the number of participants within a decision-making entity (Lambard-Mogliansky et al 2006).

<sup>10</sup>See Appendix B for the extensive form specifications of the adjusted games for all four treatments.

equal to the ones in IDT1. given the amount of bribe. This way we want to separate the true ‘Group Decision-making Effect’ (GDE) from effects stemming from the partition of the bribe between the two officials.

In Stage 2 and Stage 3. the two officials of a unit take a decision jointly. In both stages they decide independently first. If they do not come to an unambiguous decision (e.g. one subject decides for ‘reject’, the other for ‘accept’ in Stage 2), they learn about each other’s choice and decide again. If there is still no agreement, they get the opportunity to communicate with each other via a real time electronic ‘chat’ in which they can, for one minute, exchange electronic messages (see the instructions in Appendix D).

If there is still no mutual consent, the corruption-unfriendly choice is taken (‘reject’ in Stage 2 and ‘defect’ in Stage 3). This rule reflects the veto-power of officials who do not want to engage in a corrupt transaction in Stage 2 and those who do not want to reciprocate in Stage 3 (in both cases avoiding damage to the public). The idea is that an official cannot force his colleague to engage in a corrupt transaction but can try to convince him. As a consequence of the decision making rule, a corrupt transaction can only be successful if both officials choose ‘accept’ in Stage 2 and ‘cooperate’ in Stage 3.

## **IDT2**

The ‘Individual Decision-making Treatment 2’ (IDT2) differs from the treatment IDT1 only in the number of possible transactions between a certain  $B$ - $O$  pair. In this treatment we consider units of four, two type  $B$  participants and two type  $O$ . Every type  $B$  subject sends only one transfer to one of the two officials in his unit per period. This means that playing the game for ten periods makes five possible transactions per pair, producing two transactions per period and four-player-unit. The justification to run this treatment is to control for possible effects in the behaviour of subjects stemming from playing in a larger group and interacting less frequently with a particular transaction-partner.

The decisions of the participants within a unit of four yield only one independent observation.

In Stage 1. one of the (potential) bribers ( $B_1$ ) decides about her transfer  $b_1$  to one of the officials ( $O_1$ ), and the other briber ( $B_2$ ) decides about his transfer  $b_2$  to the remaining official ( $O_2$ ). All transfers are tripled and shown to the respective officials. In Stage 2 and Stage 3 each of the officials decides independently. The respective pairs change every period so that in the following period  $B_1$  decides about the size of her transfer to  $O_2$  while  $B_2$  interacts with  $O_1$ .

## TDT2

In the ‘Team Decision-making Treatment 2’ (TDT2) we form again four-player-units. Each of the two type  $B$  players sends one transfer each to the group of two officials who decide jointly in Stage 2 and Stage 3. according to the decision making process explained for TDT1. Here the transfer is split equally so that each of the officials receives only half of (the tripled amount of) the transfer chosen by the respective briber ( $3 * 0.5 * b$ ). This means that each of the type  $B$  subjects makes one decision, while each of the two officials has to decide in two separate situations per period. As a consequence, each type  $O$  subject receives two payoffs per period, while  $B$  receives only one.

Here the parameters are set in such a way that one transaction in TDT2 corresponds to two transactions in IDT2 in terms of payoffs for type  $O$  subjects. This means that  $O$ ’s individual incentives for a certain transaction are not equal to the ones in IDT2. since a certain transfer  $b$  leads to double the amount of revenue reaching a single official in IDT2 as compared to the revenue reaching each of the subjects in the role of the official in TDT2. Not only gains (tripled transfers) but also costs (4 EMU) are shared equally<sup>11</sup> between the two officials.

## 3 Hypotheses

The experimental literature on corruption (Dusek 2004. Abbink 2004. Lamsdorff and Frank 2007. 2010) provides robust evidence for substantial reciprocity and positive levels of corruption in set-ups comparable to ours<sup>12</sup> but using individual decision makers only. The findings can neither be explained by the standard self interested model (see Appendix A) nor by models of reciprocity relying on altruism or social preferences. Altruism (Andreoni and Miller 2002) fails as an explanation since the successful corrupt transaction is not only defined as decreasing the payoff of outsiders but also decreases the sum of payoffs. It is also unconvincing to use an argument based on inequity aversion or ‘MinMax’ preferences (Fehr and Schmidt 1999, 2006, Charness and Rabin 2002) since the outcome of a successful reciprocal transaction reduces the welfare of those who are potentially (externality in mode 1) or certainly (externality in mode 2) worse off than the decision-making agents in case of a

---

<sup>11</sup>The assumption of an equal split of the transfer reflects equal bargaining power and similarity of the officials. There is no reason to believe that endogenous distribution would yield a different sharing rule.

<sup>12</sup>The common feature is the trade-off between reciprocity based maximization of individual long-term payoff and a combination of myopic short term maximization of personal gains and a preference for social efficiency.

successful corrupt transaction. An equilibrium would critically depend on the assumptions on the reference group which is likely to be heterogeneous across and possibly not consistent within subjects. The clear strategic background of the situation makes an explanation of reciprocity based on intentions (Falk and Fischbacher 2006) difficult since positive transfers are unlikely to be considered as displays of kindness. More convincing explanations may be expected in the notion of ‘weak’ reciprocity based on repetition and reputation formation (Fehr and Fischbacher 2003, Kreps et al 1988).

However, even though behaviour may be (partly) explainable by a fully fledged theoretical model *ex post*, there certainly still is a lack of the ability to develop point predictions with respect to the effect of the introduction of the 4EP.

Considering stylized facts derived from experimental evidence of the corruption game (e.g. Abbink et al 2002) and studies on group decision making (e.g. Kocher and Sutter 2007), we distinguish between two main effects from the introduction of the 4EP with respect to the officials’ behaviour.

First, the introduction of the 4EP causes a bribe to be divided between two officials instead benefitting just one. Keeping  $B$ ’s behaviour (i.e. the amount of transfer  $b$ ) constant, the splitting of the bribe causes each official to receive less than what a single official would have got. We call the officials’ immediate reaction on the lower benefit from a bribe in TDT1 and TDT2. the Group Decision making treatments (GD treatments), the ‘Bribe Splitting Effect’ (BSE).

Second, we know from the experimental literature on group decision making (e.g. Chalos and Pickard 1985) that decisions made in a group can differ from individual decisions even if individual marginal incentives are the same. Experimental evidence suggests that the direction of this effect (which we call group decision-making effect: GDE) is ambiguous and depends highly on the nature of the particular situation. Kocher and Sutter (2007) show that decisions made jointly inside a group may be either more or less in line with standard game theoretical predictions (using the self interested model), depending on what kind of motive dominates the decision making process.

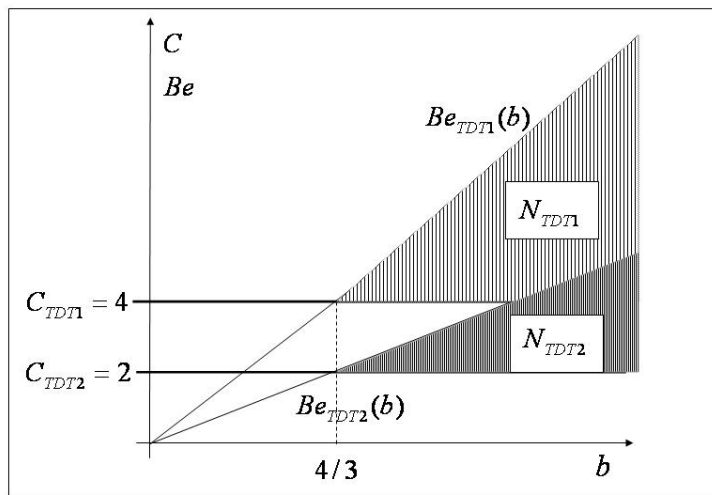
### **3.1 Bribe Splitting Effect**

From a series of experiments using comparable set-ups (e.g. Abbink 2004), we know that the probability of success increases with the level of transfer (bribe). The phenomenon relates closely to the findings of reciprocity in the gift exchange game (Fehr et al 1993)

which investigates reciprocity of workers with respect to costly effort in exchange for wage premiums. For simplicity we assume the bribe to be shared equally between the two officials, ignoring potential distributional issues.<sup>13</sup> The existence and magnitude of an effect stemming from the splitting of a bribe depends on what officials condition their behaviour on. Officials may consider the monetary benefits of the bribe only or they may include the ‘intentions’ of the briber. Leaving intentions out, the correlation between the size of the bribe and the probability of positive reciprocity may be explained by the trade-off between marginal benefits and marginal costs of engaging in corruption. Moreover, considering the repeated set-up, the minimum amount of transfer (in an early period) needs to be large enough to give  $O$  an incentive (in terms of expected benefits in the future) to make her incur personal costs with the objective to trigger  $B$ ’s reciprocity in future periods.

Figure 2 illustrates the monetary benefits and costs as a function of the level of transfer in the treatments TDT1 and TDT2. In TDT2, the benefits as well as the costs are split between the officials and hence are both half as large as in TDT1. The net monetary benefits (difference between benefits and costs) are labelled as  $N_{TDT1}$  and  $N_{TDT2}$ .

Figure 2: Monetary Costs and Benefits of corruption in TDT1 and TDT2



For  $b < \frac{4}{3}$  corrupt behaviour is not rationalizable since costs are larger than benefits.<sup>14</sup> For  $b > \frac{4}{3}$  the monetary benefit is larger in TDT1 than in TDT2 while the difference is increasing in  $b$ . If  $O$  conditions her corruptibility on the net benefit of the transaction at

<sup>13</sup>In our anonymous setting there is no reason to believe that there should be any other kind of distribution rule to be agreed by both officials.

<sup>14</sup>For  $b < \frac{4}{3}$  not even the technical costs  $c$  are covered.

all, the probability of corrupt success must be weakly larger in TDT1 than in TDT2 for all  $b > \frac{4}{3}$ .<sup>15</sup> The difference in probabilities should be even larger considering not only the monetary but also the moral costs of inflicting (monetary) harm to other members of society (i.e. other participants of the experiment in mode 1 or recipients of the donation in mode 2). Unlike the technical costs ( $c$ ), these are likely to apply to both officials at the full scale since the approval of both is needed to finalize a corrupt transaction and hence they should both be held morally accountable.

However, if subjects condition their behaviour on intentions and equilibrium outcomes alone (i.e. they consider the ‘kindness’ of  $B$ ’s decision only in a sense that it leads, given that the transaction will be successful to a certain outcome) there should not be any difference between the conditional behaviour of type  $O$  subjects in TDT1 and TDT2.

Assuming  $O$ ’s reference group to be  $B$  only (ignoring the negative externalities to the public) the payoff equalizing strategy would require the same actions across all types and treatments. Considering outcomes only, models based on inequity aversion (e.g. Fehr and Schmidt 1999) would yield the same predictions across treatments.<sup>16</sup>

**Hypothesis 1:** Holding bribe levels constant, there will be no difference in corruption levels between TDT1 and TDT2. If officials condition their reciprocal behaviour exclusively on intentions or consider equalization of payoffs only.

Hypothesis 1 will be rejected if the actual amount of bribe in a particular situation has an effect on the probability of success of a corrupt transaction and is different between TDT1 and TDT2. In this case we call the effect the Bribe Splitting Effect (BSE).

### 3.2 Group decision-making Effect

In order to measure the effect of group decision making (GDE) separated from BSE, we have to compare the behaviour of subjects deciding alone and subjects deciding within a group in situations in which all relevant decision makers face the same marginal monetary incentives. The comparison between IDT1 and TDT1 satisfies this condition.

Compare the expected number of successful corrupt transactions,  $E(N_{success})_{IDT1}$  (one official decides individually) with the expected number of successful corrupt deals,  $E(N_{success})_{IDT1}$

---

<sup>15</sup>This allows for individual heterogeneity and does not even exclude those participants who would never engage in a corrupt transaction ( $p_j(b; c) = 0$  for any value of  $b$  and  $c$ )

<sup>16</sup>This is because of the difference in the number of transactions played per period by the different types.

(two officials decide jointly) facing the same bribe  $b$ .

In IDT1. the probability of success of a corrupt transaction is  $p_i(b)$  for the deal in which official  $O_i$  is relevant. Assume that the probability of corrupt success (reciprocity) is positively dependent on the relevant bribe  $b$ . The probability of success is simply  $p_i(b)$ .

In TDT1. officials  $O_i$  and  $O_j$  decide jointly in Stage 2 and Stage 3. The group decision making process provides veto power for non-corrupt and non-reciprocal behaviour ('reject' in Stage 2 and 'defect' in Stage 3).

In this case the probability of success is  $p_i(b)p_j(b)$ .

Since  $p_i(b_1) \leq 1$  and  $p_j(b_2) \leq 1$ ,  $E(N_{success})_{IDT1} = p_i(b) \geq p_i(b)p_j(b) = E(N_{success})_{TDT1}$ .

As long as the individual behaviour of officials is independent from the decision making process (including the observation of or the belief on the behaviour of the other official), the expected number of successful corrupt transactions should be weakly greater under IDT1 than under TDT1.

**Hypothesis 2:** If decisions are completely independent from the decision making process we will observe lower (relative) numbers of successful corrupt deals in TDT1 than in IDT1.

We know from experimental research that individual decisions are far from independent when made inside a group or team. Experimental evidence suggests that the effect of group decision making (which we will call GDE in the following) tends to be more in line with the predictions of the self centred model of profit maximization when considering bargaining situations in which competition plays a relevant role, while in games representing social dilemmas, e.g. the public goods or the gift exchange game, groups may even move further away from the predictions of standard game theory (Cason and Mui 1997, Levine and Moreland 1998). Kocher and Sutter (2007) track this back to two opposing motives driving the effect of group decision making, the competitive and the profit maximizing motive. The total effect depends on which of these dominates.

The set-up of a corrupt transaction cannot easily be categorized into one of the two situations. Our set up resembles the gift exchange game in some aspects but fundamentally different in others. Where reciprocity-based cooperation increases the level of social welfare in most applications (as it usually helps to overcome a social dilemma), corrupt reciprocity decreases social welfare (and efficiency) by design.<sup>17</sup>

---

<sup>17</sup>Contrary to the models of social dilemmas it is (for the official) always (myopically) individually as well as socially optimal not to reciprocate.



### 3.2.1 Competitive Motive

The pure fact of group membership may cause a shift in individual preferences towards a decision that reflects higher awareness of competition with other groups or individuals (minimal group paradigm, Tajfel and Turner 1986). In our situation this may push groups to behave more in line with the predictions of standard game theory. Members of a group might follow strategies that increase the difference of pay-off levels between in- and out-group members. Especially under mode 1 (fellow subjects are hit by negative external effect), this may help cooperation within and hinder cooperation across groups.

An important question, which cannot easily be answered in our setting is, who exactly type  $O$  subjects consider as their group-members. These could be the fellow official (who a subject forms a group with), but it could also be the unit of all participants involved in the transaction, including the type  $B$  player(s).

Being primarily interested in behavioural differences between individual decision making treatments (ID treatments) and group decision making treatments (GD treatments), the in-group-effect within the transaction unit (including respective  $B$  and  $O$  participants) should be considered as irrelevant, since it should be equally present across treatments. What differs between ID and GD treatments is the additional in-group-effect between jointly deciding officials in the GD treatments. The creation of a sub-group by letting officials decide jointly may result in more competitive behaviour towards their type- $B$  transaction (unit) partners which could result in myopic profit maximization at the cost of corrupt reciprocity. We denote this motivation as the competitive motive (CM).

### 3.2.2 Profit Maximizing Motive

The *Profit Maximizing Motive* (PMM) causes groups to make decisions that yield larger profits (in the long run) while, if necessary, shifting behaviour even further away from equilibria predicted by standard theory (Kocher and Sutter 2007).

Despite being inefficient on a social level, a successful corrupt transaction yields the largest profits for the transaction partners.

Groups may be more capable of suppressing short-sighted impulses of behaviour which may maximize myopic profits but ultimately decreases total individual profits of all transaction partners. This behaviour includes free-riding or defecting in public dilemmas (e.g. public good game using voluntary contribution mechanism) and failing to foresee the breakdown

of future cooperation (reciprocal relationships). The persuasive argument theory (PAT, see Pruitt 1971, Bishop and Myers 1974, Burnstein et al. 1973) predicts that groups are more successful in finding strategies that maximize their members' long term payoff. Explorative and knowledge capacities in groups (containing more than one individual) are expected to be greater than those of a single individual.

Chalos and Pickard (1985) proclaim that groups are better in processing information load. In games where profit maximizing strategies are as complicated as in the repeated corruption game, we expect groups to develop and follow more successful strategies than individuals with respect to maximizing their members' monetary payoffs when we assume groups and individuals to exhibit equal preferences with respect to the trade-off between individual and social welfare maximizing.

**Hypothesis 3:** If the group decision making process is dominated by the CM, outcomes will be closer to the game theoretical predictions in TDT1 than in IDT1. If the PMM dominates group decision making, groups will produce higher levels of corruption by following strategies that are more successful in maximizing their members' individual payoffs.<sup>18</sup>

### 3.3 *B*'s behaviour

The introduction of the 4EP may not only affect the behaviour of the officials but also that of the bribers. The direction of the effect depends entirely on the beliefs about the (effects on the) behaviour of the official(s).<sup>19</sup>

Even if we assume very restrictive belief structures it is hard to formulate consistent hypotheses.

If subjects of type *B* expect groups of officials to be less likely to reciprocate (given a certain amount of bribe) than individual officials, *B*'s reaction can (still) go into both directions.

On the one hand, bribers who want to initiate a corrupt transaction might be discouraged by their anticipation of a higher probability of failures and therefore choose '0'-transfers more often in the GD treatments.

On the other hand, there might be an increase in the bribe level in the GD treatments, com-

---

<sup>18</sup>Note that the comparison between IDT1 and TDT1 shuts off any potential effect arising from the splitting of the bribe between the two officials deciding jointly in a group since marginal incentives for all subjects are the same in both treatments.

<sup>19</sup>Relying on the assumptions of standard game theory (within the self interested model), we do not expect bribers to have different beliefs about the behaviour of groups and individuals, since in all treatments the Subgame-perfect Nash Equilibrium (Appendix A) is unique and predicts neither positive transfer levels nor positive reciprocation.

ing from bribers who anticipate that groups of officials are more demanding to be ‘convinced’ to act in a reciprocal way than individuals are.

Therefore the total effect is ambiguous and depends on which of the effects dominates the decisions.

The argument for a belief structure that assumes higher reciprocity (for any bribe level) within groups of officials leads to similarly ambiguous predictions.

In contrast to the inconclusive predictions on  $B$ ’s expected reaction on the anticipation of the GDE, we may be able to predict the direction of the effect stemming from  $B$ ’s anticipation of the BSE which can be quantified by comparing average bribe levels between TDT1 and TDT2. If bribers anticipate the BSE (correctly) they will probably send larger transfers in order to compensate the splitting of the bribe.<sup>20</sup>

**Hypothesis 4:** The bribe level (and distribution) will be different in TDT1 and TDT2 if type  $B$  subjects anticipate officials to show a BSE and react appropriately.

### 3.4 Gender Effects

An especially important finding in the empirical literature on corruption is the relationship between gender and corruption (Lambsdorff et al and Frank 2010). Dollar et al (2001) and Swamy et al. (2001) find that female participation in market transactions leads to lower levels of corruption. Sung (2003) however advocates that the findings might be at least partly due to inconsistencies related to omitted variable biases.<sup>21</sup>

Avoiding these problems, Rivas (2007) as well as Lambsdorff and Frank (2007) use controlled laboratory experiments to find that women tend to reciprocate less often with the corrupt partner in the role of the official even if risk aversion, a potential source of behavioural gender difference, can be ruled out as a driving force (Schubert et al 1999).<sup>22</sup>

---

<sup>20</sup>Note again that such a reaction could not be explained by models of inequality aversion since a strategy that aims at equalizing payoffs would not proclaim different levels of transfers across these treatments.

<sup>21</sup>The real but unobserved driver could be the quality and level of development of institutions, affecting both, women labour market participation and corruption. Assuming a positive correlation between institutional quality and women (labour market) participation, we would expect an over-estimation of the predicted effect. Furthermore women might self-select into sectors where corruption is less rampant. This could be motivated by a vector of non-measurable variables presenting even more complex problems of misspecification.

<sup>22</sup>Note that experiments on corruption that model the individual cost of reciprocation by adding a lottery including a large loss of payoff with small probability as a fourth stage of the corruption game (instead

This behavioural difference across genders may be explained by the fact that female participants are less inclined to anticipate negative reciprocity.

The findings lead to the conclusion that the presence of women can destabilize trust and reciprocity-backed stability of the corrupt transaction between  $B$  and  $O$ . It provides a strong argument for policies aiming at increasing women participation in corruption-sensitive sectors within public procurement. Lamsdorff and Frank (2010) show that the gender effect is especially strong when  $B$  has a direct opportunity of negative reciprocity, i.e. costly punishment of defective behaviour (by the tool of whistle blowing). In our set-up  $B$ 's only opportunity of negative reciprocity is by reducing bribes in future periods.

Relying on the results of corruption experiments (Lamsdorff and Frank 2010) using comparable set-ups we do not expect a strong gender effect in  $B$ 's behaviour as long as  $O$ 's gender is unknown.

The analysis is more complicated when considering the gender effects with officials. For the individual decision making (ID) treatments, we expect a gender effect only if the female lack of anticipation of negative reciprocity extends to cross-period reciprocity (i.e. anticipating lower transfers in periods following a defected transfer situation).

In order to control for gender effects within officials in the GD treatments, we need to distinguish between pure female groups (both officials are women), mixed groups and pure male groups.

Even if we assume a gender effect within officials' decision making, the interaction of a male and a female official, when deciding jointly, is unclear.

**Hypothesis 5:** If the gender effect is present in our repeated set-up, female officials in the ID treatments and pure female groups of officials in GD treatments will produce lower levels of successful corrupt transactions than male officials in ID treatments and pure male (and mixed groups) in the GD treatments.

### 3.5 Total effect

In order to evaluate the usefulness of the implementation of the 4EP, the total effect, i.e. the combination of the BSE and the GDE is relevant. This effect can be directly measured by the comparison of IDT1 (or IDT2) and TDT2.

---

of modelling it as a certainty equivalent) cannot distinguish corruption-specific gender effects from gender effects that are caused through differences in individual risk aversion without controlling for individual risk attitudes separately.

The institution of the 4EP can only help to reduce the level of corrupt activity in two situations. The 4EP will certainly reduce corruption if the BSE and the GDE are both positive. It will also reduce corruption if the GDE is dominated by the competitive motive (and therefore negative) but overcompensates by the BSE. Only if the GDE is negative and overcompensates the BSE, the introduction of the 4EP is counterproductive even without considering the costs of the installation of such an institution.

**Hypothesis 6:** If the GDE is negative (PMM dominates CM) and dominates the BSE, the average rate of corrupt success will be greater in TDT2 than in either IDT1 or IDT2.

## 4 Procedure

All 8 sessions (two sessions for each treatment) were programmed and conducted at the experimental laboratory MELESSA at the University of Munich. It used the program Z-Tree (Fischbacher 2007) and the organizational software Orsee (Greiner 2004). Each session contained 24 subjects.

Subjects were randomly assigned a type, ( $B$  or  $O$ ) and randomly allocated into groups of two in IDT1, into groups of three (one type  $B$  and two type  $O$  participants) in TDT1 and groups of four (two type  $B$  and two type  $O$  participants each) in IDT2 and TDT2.

In all treatments group members stayed together in their groups for all 10 periods (partner design) where full anonymity was ensured.<sup>23</sup>

Every period was paid where 1 EMU was worth 5 eurocents. Payoffs were summed up over all 10 periods and paid out in private at the end of the experiment. The whole experiment took less than 90 minutes. The instructions were kept completely neutral, avoiding any language indicating the subject of research in order to concentrate on the specific features of the model and minimize the difference between the instructions of the treatments. Abbink et al (2006) show that framing has no significant effect on behaviour in the corruption game. All this was common knowledge to all participants. Understanding of the (rather complicated) set up was made certain by partly reading out the instructions, answering questions in private and checking of several control questions.

In the end of the experiment subjects filled out a questionnaire including demographic information. Payoffs lay between 4 Euros and 25 Euros excluding a show up fee of 4 Euros which is standard to experiments at MELESSA. Average earnings amounted to 14.37 Euros.

---

<sup>23</sup>Note that the interaction of officials was conducted via a chat which did not allow for any form of identification, see the Appendix D for the full instructions of TDT2.

A total of 249.10 Euros was paid out as a donation to the organization ‘Doctors without Borders’ as a result of the decisions made by the participants in the treatments where we chose mode 2 as a model of the negative externality (one session in IDT2 and one session in TDT2).

## 5 Results and Interpretation

As noted in Section 2. the negative external effect has been modelled in two different ways. According to predictions using the standard self interested model, there should be no effect of either of the two models and hence no difference between the set-ups, neither in the behaviour of the official (or the group of officials) nor in that of the briber. In our experiment, the difference might play a role for the size and even direction of the GDE. In mode 1. the negative external effect hits four randomly chosen participants of the same session who can themselves make decisions that result in negative payments to other participants. This interaction between groups may produce an unwanted additional in-group effect and destroy the idea of the external effect as an unreciprocated (and unreciprocateable) reduction in the payoff of unrelated third parties. In particular subjects may justify their own corruption by their belief of others’ corruption. Corrupt behaviour might even be considered as a payoff equalizing equilibrium (‘super-game’).

In order to rule out effects stemming from these considerations we applied mode 2 in one session for IDT2 and one session for TDT2. In mode 2. the ‘super-game’ problem is eliminated by modelling the negative externality as a reduction of a (fixed) donation to ‘Doctors without Borders’.<sup>24</sup>

Checking for differences we compare all relevant variables, i.e. the average total transfer-level, the transfer levels after success and failure of a corrupt deal (measure of the client’s reciprocity), the average relative number of successful deals, the percentage of rejected bribes and the percentage of zero-value transfers between observations of the two modes and find no significant difference, taking averages over all 10 periods for each (relevant) subject and applying (pair-wise) Mann Whitney U-tests<sup>25</sup> on group levels ( $p \geq 0.36$ ;  $N = 12$ ). We take this as sufficient evidence for the assumption that the design of the externality (mode) does not have any impact on the outcomes of the relevant decisions and therefore pool the data from these sessions in the respective treatments for the entire analysis.

---

<sup>24</sup>We chose this organization to be able to compare our results to those obtained in Lambsdorff and Frank (2007) who use donations to this organization to model the negative externality.

<sup>25</sup>Unless stated otherwise (exact, highest or lowest)  $p$ -values and numbers of observations (N) apply to the Mann Whitney U-test.

The main objective of the experiment is to evaluate the introduction of the 4EP by comparing the main performance variables in IDT2 and in TDT2. This could have been problematic since these treatments do not only differ in the decision making process and splitting of the bribe (introduction of the 4EP), the officials also face unequal numbers of transactions per period. Where in TDT2 an official interacts (together with the other official) with each of the two bribers (in his unit of four) once in a period, an official in IDT2 only interacts once in every two periods with a particular briber (in his unit of four). This makes a large difference with respect to the ‘horizon’ of bilateral (multilateral) repetition. Considering repetition as a major determinant of reciprocal behaviour there may be differences in corruption behaviour not related to any of the effects discussed in Section 3 (GDE and BSE). A comparison between the outcomes of IDT2 and TDT2 may still be interpreted as capturing the total effect of the introduction of the 4EP when considering that it may also lead to an increase in the frequency of multilateral transactions, which is not unreasonable when the available number of officials is held fixed between the settings.

To check whether there is an effect of the amount of repetition we compare the outcomes of IDT1 and IDT2. Although we do not find any significant differences in the main variables (U-tests;  $p > 0.23$ ;  $N=24$ ), we do not pool the data from these treatments, but apply separate tests. For reasons of clarity we pool the data only for graphical illustrations.

## 5.1 Descriptive results

Table 1 shows the values of the four main performance variables in all four treatments. Variable *Corrupt success* depicts the average share of successful deals per unit ( $\frac{N_{success}}{N_{total}}$ ).<sup>26</sup> *Payoff* presents the average payoff level (in Euros) after the reduction of the negative externality (final payoff) per subject. *Neg. Externality* describes the level of the (relevant) negative externality in Euros. To be able to accurately compare payoff levels between all treatments, we subtract the relevant share of the reduction of the donation from the actual payoff in the externality mode 2. *Transfer level* measures the average amount of bribe (in EMU) transferred by type *B* participants per period.

---

<sup>26</sup>Note that units contain two, three or four individuals which makes a comparison between absolute levels of corrupt success inconclusive.

Table 1: **Performance variables**

Treatment	<i>Corrupt success</i>		<i>Payoff</i>		<i>Neg. Externality</i>		<i>Transfer level</i>	
	mean	std.dev.	mean	std.dev.	mean	std.dev.	mean	std.dev.
IDT1	0.25	0.09	13.13	3.92	2.90	1.01	3.18	1.52
IDT2	0.25	0.06	12.72	3.78	3.49	2.12	3.16	1.25
TDT1	0.47	0.07	7.23	4.26	5.13	2.91	3.58	1.15
TDT2	0.42	0.05	9.40	4.05	5.38	1.97	4.15	1.24

All means are calculated as averages across periods and (relevant) participants of the respective treatment.

### 5.1.1 Corruption levels

Comparing corrupt success rates (*Corrupt success*) between treatments, we can identify a substantial BSE and a GDE. The (5%) difference in corrupt success rates between TDT1 and TDT2 is significant ( $p = 0.08$ ;  $N = 24$ ) proving the existence of a (relatively weak) bribe splitting effect. We reject **Hypothesis 1**.

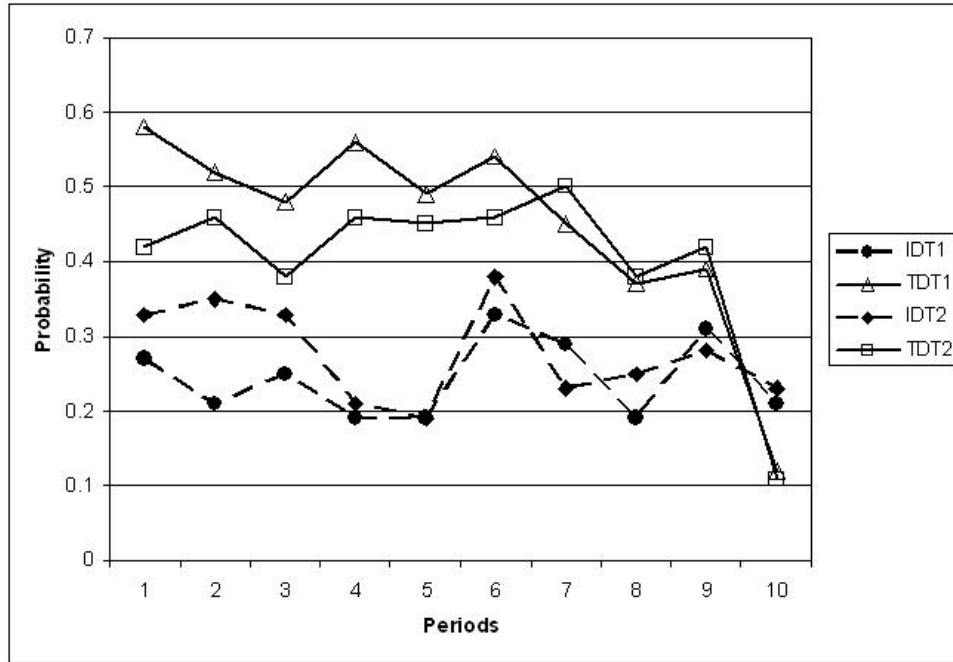
Moreover we find a large group decisionmaking effect (**Hypothesis 2**) which amounts to  $-22\%$  ( $p = 0.03$ ;  $N = 36$ ) and strongly suggests the dominance of the profit maximizing motive (**Hypothesis 3**).

Moreover, the large and significant difference (17%) between IDT1/IDT2 and TDT2 ( $p = 0.002/0.04$ ;  $N = 36/24$ ) indicates a negative total effect of the introduction of the 4EP even if we control for the difference in the number of repetitions (between IDT1 and TDT2) confirming **Hypothesis 6** (PMM dominates CM and is greater than BSE).

Figure 3 demonstrates the differences in the dynamic development of success-probabilities over all 10 periods between the ID and GD treatments. Subjects in the GD-treatments start with very high levels of successful corruption which decline gradually, showing a large ‘end-game effect’ (end of repetition in the 10<sup>th</sup> period) while corruption levels stay relatively stable (at a lower level) in both ID treatments. We interpret this as a first hint at the hypothesis of groups following systematically different strategies compared to individuals.



Figure 3: Success-probabilities over Periods



### 5.1.2 Payoff and externalities

The average net<sup>27</sup> payoff is more meaningful in terms of a welfare comparison between the treatments than the levels of corruption. Figure 4 shows the average total payoff level (*Payoff*) per subject in comparison with the average negative external effects (caused by the average subject, *Externality*) for all four treatments.

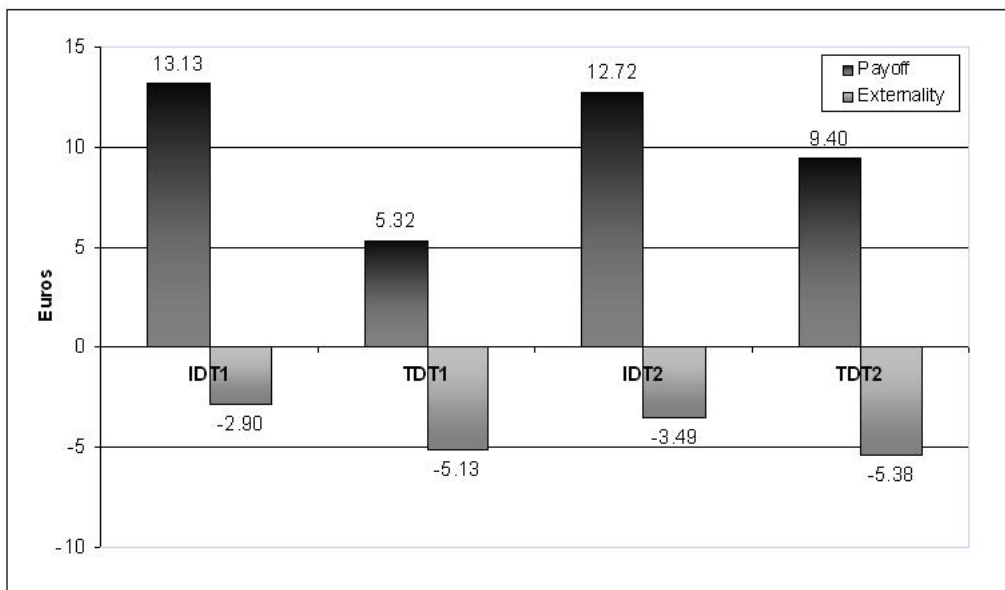
The bilateral differences can be explained through two effects related to the transfer level and corruptibility.

First, the level of transfer increases the payoff independent from the outcome since bribes are tripled (assumed marginal utility from money between  $O$  and  $B$ ) and hence enhance efficiency and profits. Second, the level of corrupt transactions (which is positively correlated with the transfer level, Section 5.2) reduces payoff levels through the impact of its negative externality.

The relatively large (but marginally insignificant) difference in payoff levels between TDT1 and TDT2 (2.17 Euros;  $p = 0.15$ ;  $N = 24$ ) is due to lower levels of transfers and higher levels of successful corruption. The relatively low difference in payoff levels between both ID treatments and TDT2 can be explained by the negative difference in bribe levels and the positive difference in the level of corrupt success. While we can only show a significant

<sup>27</sup>As net payoff we define the sum of payoffs after subtraction of the externality caused by others' decisions.

Figure 4: Payoff levels and levels of externalities



difference in payoff levels between IDT1 and TDT2 ( $p = 0.07$ ;  $N = 36$ ) and not between IDT2 and TDT2 ( $p = 0.15$ ;  $N = 24$ ), all differences between ID treatment and TDT1 are significant on the 1%-level. Taking cumulated payoffs as a measure for welfare suggests a negative total effect of the introduction of the 4EP. To consolidate this result we can report huge differences in the distribution of payoffs between the treatments. Table 2 depicts the values of the Gini coefficient<sup>28</sup> in all four treatments.

Table 2: Gini Coefficients

	IDT1	IDT2	TDT1	TDT2
Gini	0.14	0.15	0.18	0.16

The values of the Gini coefficients of sessions in mode 2 are ‘adjusted’ for the negative externality.

In both GD treatments some units accumulate large payoffs by sending high transfers and profiting from positive reciprocity at the expense of other participants (mode 1) whose units do not engage in corruption (as often) but are hit by the negative external effect just the same. This leads to more inequality in payoffs.

<sup>28</sup>The Gini-coefficient measures the distribution of wealth within a certain population on a scale between 0 and 1 where a value of 0 corresponds to complete equality and a value of 1 signifies complete inequality.

## Payoff maximizing strategies of *B*

Assuming that subjects aim at maximizing personal profits, type *B* subjects' strategies vary systematically in their effectiveness between the treatments. We use a simple linear panel regression (random effects) to derive the profitability of bribing.<sup>29</sup> To account for dynamic aspects and potentially decreasing (or increasing) profitability of transfers we use the following specification for 88 (all) participants of type *B*. We cluster errors on the unit level to take possible dependence of behaviour between type *B* participants of the same unit into account.

$$PP_{it} = \beta_0 + \beta_1 b_{it} + \beta_2 b_{it}^2 + \gamma D_i + \delta D_i * b_{it} + \zeta D_i * b_{it}^2 + \epsilon_{it} \quad (M1)$$

The dependent variable *PP* signifies the individual payoff (in EMU) excluding subtractions from the externality caused by subjects outside a subject's unit. Independent variable *b* indicates the transfer paid in the respective period. The quadratic variable  $b^2$  captures a potential non-linear effect of the transfer on period payoffs (decreasing marginal payoffs). Vector *D* stands for the treatment Dummy variables  $D_{IDT2}$ ,  $D_{TDT1}$  and  $D_{TDT2}$ . Accordingly, vectors  $D*b$  and  $D*b^2$  contain interaction effects for the treatments TDT2, TDT1 and TDT2.

We find (see Table 3 in Appendix B for the full list of estimated coefficients) that the individual payoff per period is strictly increasing in the level of transfer sent to the (group) of official(s) in all treatments.

There is no significant evidence for a decreasing profitability of bribing ( $\beta_3$ ;  $p = 0.26$ ) in neither of the treatments. The main finding is that bribing is far more profitable in the GD than in the ID treatments. On average an additional unit of transfer increases the payoff per period by 1.89 EMU in IDT1 and 2.02 in IDT2 (difference is not significant  $p = 0.28$ ). By contrast, the marginal effect is 3.94 EMU per unit in TDT1 and 2.63 EMU in TDT2. The differences between these marginal effects and the ones in the ID treatments are all highly significant (F-test;  $p < 0.01$ ).

## Payoff maximizing strategies of *O*

The differences in the individual profitability from engaging in corrupt activities between the treatments are not as large for the officials.

We estimate a simple OLS regression, measuring the marginal effect of *N*, the number of

---

<sup>29</sup>Using pooled OLS (with and without period dummies) did not yield qualitatively different results.

successful corrupt transactions (number of choices in which the (group of) official(s) cooperated) on *Payoff*, the total payoff of each (group of) official(s).<sup>30</sup>

$$Payoff_i = \beta_0 + \beta_1 N_i + \gamma D_i + \delta D_i * b_i + \epsilon_i \quad (M2)$$

Vector  $D$  and  $D * b$  have the same interpretations as in Model M1. The results of the regression are reported in Table 4 of Appendix B. We find a strong positive effect of the number of successful transactions on the total payoff in all treatments. The profitability of being corrupt is significantly higher in both GD treatments than in the ID treatments (four F-tests:  $p \leq 0.004$ ). While officials in the ID treatments earn on average only 16 EMU more for an additional successful corrupt transaction, the rate is at 19 (TDT1) and 21 (TDT2) considerably higher in the GD treatments.

## Altruism

To accept a bribe instead of rejecting it in Stage 2 (of the game), means to keep the benefit (of the tripled transfer) for oneself instead of sharing it with the public. Moreover, the corruption game is designed in such a way that for any given level of transfer, the socially optimal decision (maximizing the sum of payoffs) is to reject the bribe in Stage 2.

The number of non-zero<sup>31</sup> rejections of bribes in Stage 2 is extremely low in all four treatments.

In only 6.75% of possible cases (i.e. if  $b > 0$ ) (groups of) type  $O$  subjects reject a non-zero bribe in IDT1 compared to 6.22% in TDT1. 5.40% in IDT2 and 5.44% in TDT2. The results indicate that displays of altruism towards the public at a high personal cost ( $2b$ ) is rare. Moreover we do not find a significant difference across treatments (U-test;  $p \leq 0.47$ ;  $N \geq 24$ ).

### 5.1.3 Transfer levels

Average bribe levels (including 0-transfers) are substantially (and significantly) larger than 0 for all treatments and almost identical within the ID treatments (3.18 and 3.16 EMU). Transfers are only slightly (and insignificantly:  $p = 0.35$ ,  $N = 36$ ) larger in TDT1

---

<sup>30</sup>Since officials within a unit in the GD treatments decide jointly they receive the same payoffs and are treated as a single observation. Officials being in the same unit but deciding independently (IDT1) are treated as individual observations but we cluster their standard errors in the regression.

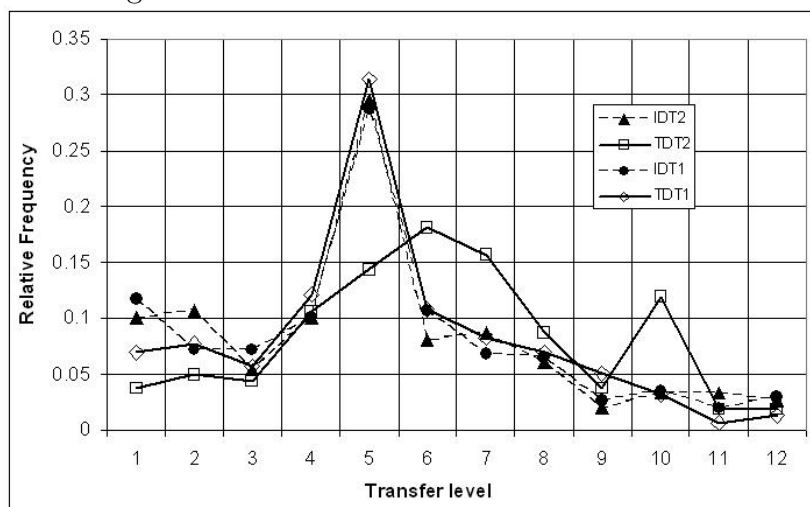
<sup>31</sup>Accepting a 0-level bribe should be (weakly) dominated by the option ‘reject’ since it inflicts damage to the public with no personal gain. Therefore an interpretation as altruistic behaviour cannot be justified and this occurs only twice as an outcome in the experiment.

(3.58), while, at 4.15 EMU, the average transfer level in TDT2 is significantly ( $p \leq 0.04$ ;  $N \geq 24$ ) larger than the ones in any of the other treatments.

The large difference in bribe levels between TDT1 and TDT2 suggests that bribers anticipate different behaviour from officials and react accordingly. Taking into account that success levels in corruption are significantly lower in TDT2 than in TDT1 despite the positive difference in transfer levels, we conclude, assuming realistic beliefs, that bribers anticipate the BSE and ‘react’ by trying to ‘convince’ officials by transferring larger bribes (**Hypothesis 4**).

The distribution of the size of transfers reveals even more information about  $B$ ’s behaviour. Figure 5 shows the relative frequency of transfer levels in all four treatments. Transfers are almost identically distributed in IDT1. TDT1 and IDT2. There are few transfers for low ( $b < 4$  EMU) and high ( $b > 8$  EMU) with a very strong mode at  $b = 5$ . This (particular observation) may be explained by subjects acting according to inequity aversion (Fehr and Schmidt 1999). The strategy ( $b = 5$  EMU; ‘accept’; ‘cooperate’) leads to equal payoffs for  $B$  and  $O$  within a unit in all four treatments.<sup>32</sup>

Figure 5: Relative distribution of transfer levels



The distribution in TDT2 depicts a significantly (Kolmogorov-Smirnov test; exact  $p < 0.001$  for comparison between TDT2 and all three other treatments for observations of strictly positive bribes:  $N \geq 382$  as well as for all observations:  $N = 480$ ).

<sup>32</sup>For the results to be explained by social preferences we either need to assume that  $B$  and  $O$ ’s reference group excludes the public (other participants in mode 1 or recipients of donations from ‘Doctors without Borders’ in mode 2), or assume a certain structure of beliefs on the behaviour of the other units (only valid for mode 1).

In TDT2 probability mass is shifted towards the higher end of transfer levels. The second mode at  $b = 10$  EMU can be explained by a reaction to the bribe splitting. Within a certain situation between a particular briber and two officials, the strategy ( $b = 10$  EMU; ‘accept’; ‘cooperate’) leads to equalized payoffs. However, since the total payoff for a certain period consists of two payments for each of the officials while the briber only receives one, this strategy does not equalize outcomes with respect to the total period payoff. It is common knowledge (and was made explicitly clear with the help of several control questions in the instructions) that it is the strategy ( $b = 5$  EMU; ‘accept’; ‘cooperate’) which yields equal outcomes (in expectations) for all transaction partners, just as in the other treatments.

A possible explanation for second peak may be that bribers see the necessity to compensate officials for their non-monetary costs in engaging in a corrupt transaction. While the monetary benefits as well as the monetary costs are split in TDT2 (compared to TDT1) we may interpret the higher levels of transfer levels in TDT2 as a premium compensation for the ‘moral’ costs of causing damage to the public, which applies to both participants, since they both have full responsibility<sup>33</sup> for the corrupt outcome.

## 5.2 Conditional reciprocity

### 5.2.1 Conditional reciprocity of $O$

Throughout the experimental literature on trust and reciprocity in general (Fehr et al 1993) and on the corruption game in particular (Abbink et al 2002, Abbink 2004, Lamsdorff and Frank 2010), the scale of reciprocation has been found to depend critically on the first mover’s behaviour (in our case the level of  $b$ ). In order to explain differences in the level of corruption between our treatments we need to control for the level of transfer. Figure 6 shows the probability of a successful corrupt deal (in %) for any positive bribe level (1-12 EMU) in all four treatments.

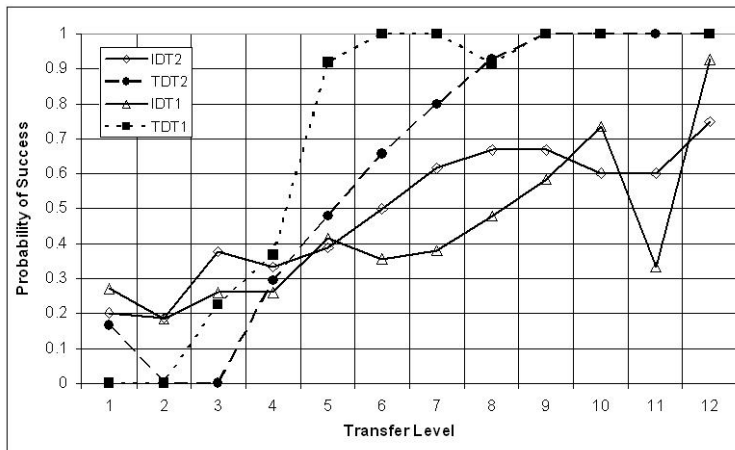
In all treatments success rates increase with the level of  $b$  (Spearman rank correlation coefficients:  $\rho \geq 85\%$ ;  $p < 0.01$  for all treatments). This is not surprising since the cost of corruption is fixed and future gains from successful reciprocity increase with the size of the bribe, assuming that bribe levels are positively correlated across periods.

While the relationship between success and transfer levels seems almost linear in IDT1 and IDT2. we observe a pattern for the GD treatments. In these treatments the conditional probability is substantially higher for large transfers ( $b > 5$  EMU) and slightly lower for small

---

<sup>33</sup>See Appendix A for an explanation of full responsibility.

Figure 6: Success rates conditional on Transfer



ones ( $b < 5$  EMU). We find significant ( $p < 0.1$ ;  $N \geq 12$ ) positive differences only for  $b > 5$  (considering transfer levels separately). The negative differences for  $b < 5$  are not significant.

Parametrically, the most straight-forward way to quantify the differences in the probability of a successful corrupt deal conditional on the relevant transfer between the treatments is to use a linear panel regression (random effects) controlling for clustered standard errors on unit level. As a robustness check we also apply pooled OLS with and without dummy variables for periods.

Since we are primarily interested in the causal relationship between the level of transfer ( $b$ ) and the success levels ( $SC$ ), we do not distinguish between a corrupt deal that failed in Stage 2 or in Stage 3. Treating the decisions ‘reject’ and ‘defect’ equally with respect to the outcome of a corrupt deal (success or failure), we do not have to take the selection process of reaching Stage 3 into account.<sup>34</sup>

We use the following specifications for the linear probability model:

$$\text{Prob}(SC_{it} = 1 | \beta X) = \beta_0 + \beta_1 b_{it} + \gamma D_i + \delta D_i * b_{it} + \theta Z_i + \epsilon_{it} \quad (\text{M3})$$

Again, vectors  $D$  and  $D * b$  stand for treatment dummies and interaction terms of treatment dummies with the transfer  $b$  as in Model M1. Vector  $Z$  contains individual demographic characteristics (e.g. age, gender<sup>35</sup>, an interaction term between gender and the level of transfer, etc.) obtained from the questionnaire. Since we do not find any significant effects with

<sup>34</sup>Treating the outcomes of ‘reject’ and ‘defect’ differently would require a Heckman-selection process explaining the selection of cases reaching Stage 3 (Heckman 1978).

<sup>35</sup>For officials in the GD treatments we use a dummy for ‘all-female’ groups and do not distinguish between ‘all-male’ and mixed groups.

any of these characteristics we do not report them in the regression output (Table 3).<sup>36</sup>

Table 3: **Output of the Linear Probability Model M3**

	Dependent variable: $SC$		
	Coefficients	Standard errors	
Constant	0.058***	0.021	
$D_{IDT2}$	0.098***	0.035	Pseudo $R^2 = 0.31$
$D_{TDT1}$	-0.034**	0.016	Number of subjects: 96
$D_{TDT2}$	-0.072***	0.023	Number of clusters: 64
$b$	0.059***	0.013	Number of periods: 10
$D_{IDT2} * b$	-0.020	0.017	
$D_{TDT1} * b$	0.052***	0.016	
$D_{TDT2} * b$	0.046***	0.013	

\*\*\* denotes significance at the 1%-level, \*\* denotes significance at the 5%-level, and \* denotes significance at the 10%-level.

In all treatments, we find that an additional unit in transfer increases the probability of the corrupt success significantly (1%-level). The effect is significantly stronger in both GD than in the ID treatments ( $\beta_{3/4} > 0$ , t- and F-tests;  $p < 0.01$ ). There is no indication for the approval of the hypothesis of women being less reciprocal in a corrupt transaction (than men) when we consider observations from all four treatments.

The behaviour of female officials (in ID treatments) and ‘all-female’ groups of officials (in GD treatments) is not found to be different from the one of their male (‘all-male’ and mixed group-) counterparts with respect to corrupt reciprocity. Female participants do not seem to behave more in line with the predictions of the standard self-interested model (**Hypothesis 5**). An explanation may be the absence of the possibility of direct punishment of defective behaviour by the bribers (e.g. negative costly retaliation) which is believed to be the main determinant of the gender effect found in Lamsdorff and Frank (2010).<sup>37</sup>

The non-linear function of success-probabilities on transfer levels (Figure 6) is best estimated by a simple maximum likelihood model. To account for differences in the marginal

<sup>36</sup>As a robustness check we ran the panel regression with a series of specifications, including a regression excluding  $Z$  and a set of pooled OLS regressions including dummy variables for periods. None of these specifications yield results qualitatively different from those reported in Table 3.

<sup>37</sup>See Chapter 3 for a comprehensive discussion of the gender effect.



effect of an additional unit in transfer on the success probability across transfer levels we run the following Probit model in its panel version (random effects).<sup>38</sup> We use the same set of independent variables and repeat all robustness checks (pooled version etc.) as in the linear probability model (M3).

$$\text{Prob}(SC_{it} = 1|\beta X) = \phi(\beta_0 + \beta_1 b_{it} + \gamma D_i + \delta D_i * b_{it} + \theta Z_i) \quad (\text{M4})$$

As expected, qualitative results (direction and significance of the evaluated marginal effects at the mean  $\bar{b} = 3.46$ ) do not change compared to the results from the linear probability model (Table 3). We find that the differences in the marginal effectiveness of bribing (considering the marginal effect of an additional unit of transfer on the probability of success, conditional on the level of transfer) between the ID and TD treatments are even larger than the differences estimated in the linear probability model for transfer levels between 5 and 9 EMU.

Table 6 in Appendix B reports marginal effects of the relevant<sup>39</sup> variables as well as predicted conditional probabilities of success. The non-linear probability model shows that these are lower in the TD than in the ID treatments for low transfers ( $b < 3$ ) while they are (significantly) larger for  $b \geq 4$ .<sup>40</sup>

The pattern shown in Figure 5 and quantified in the probit estimation (higher curvature of the probabilistic cumulative distribution function) may be explained by differences in the distinction of the strategies between groups and individuals.

On the one hand, groups of officials (GD treatments) seem to ‘defect’ (or ‘reject’) more often in the case of low transfer levels. On the other hand they seem to be more likely to reward high transfers than their individual counterparts (ID treatments). We interpret this as strategic signals of unwillingness to return the corrupt favour in less profitable transactions (aiming at inducing a higher transfer in the following period) and signals of willingness to reciprocate for high transfers (aiming at receiving further high transfer in future period in exchange for cooperation). This strategy aims at the extraction of a maximum amount of cumulative bribes. In all treatments, most non-zero transfers over all ten periods (between 79% in IDT1 and 84% in TDT2) fall into the interval for which the probability of success is

---

<sup>38</sup>See Pereira et al (2006) and Gneezy and List (2006) for examples of the use of a panel version of maximum likelihood models in comparable settings (i.e. the gift exchange game).

<sup>39</sup>Again we do not report any coefficients that are not significant, e.g. a dummy variable for gender.

<sup>40</sup>We confirm (part of) this result by finding that the average success levels conditional on the transfer level are significantly larger in (both) GD treatments than in (both) ID treatments for  $6 \leq b \leq 11$  and smaller for  $b = 2$ , using pairwise U-tests ( $p \leq 0.08$ ;  $N \geq 22$ ). Differences at  $b = \{0; 1; 3; 4; 5; 12\}$  are not significant at any relevant level ( $p > 0.12$ ;  $N \geq 7$ ).

significantly larger in the GD than in the ID treatments. Hence the strategies followed by groups seem to be more successful in the sense of higher reciprocal stability between briber and official than the strategies applied by individuals. The argument to this finding can be underlined by the content analysis of the chat-messages sent during the decision making process between officials within a group in the GD-treatments. We interpret this as strong evidence for the dominance of the profit-maximizing motive (**Hypothesis 3**).

### 5.2.2 Conditional reciprocity of $B$

The stabilization of the corrupt transaction does not only depend on the reciprocal behaviour of the official but also on that of the briber. In order to test for conditional reciprocity on the side of the briber, we compare the average size of transfer between periods that follow a successful corrupt transaction and those that follow a failed<sup>41</sup> transaction.

We apply a method introduced by Abbink (2004) that requires calculating the difference in average bribe transfers after a successful and after an unsuccessful deal (ignoring the history of events up to this point in time) for every type  $B$  participant. The resulting variable is called  $R$  ( $R = \frac{\sum b_s}{N_s} - \frac{\sum b_f}{N_f}$ ).

Replicating the qualitative results reported by Abbink (2004), briber's reciprocity is strictly positive for an average of 83% (79.2% in IDT1. 83.3% in IDT2. 87.5% in TDT1 and 83.3% in TDT2) of (the total of 88) type  $B$  participants in all treatments. This indicates a strong positive correlation between former corrupt success and the magnitude of the bribe in the next period.<sup>42</sup>

Additional to the simple  $R$ -values we also calculate an adjusted  $R$ - value which considers only observations of type  $B$  subjects who exhibit at least one case of corrupt success and at least one case of corrupt failure.<sup>43</sup> We can reject the hypothesis that  $R$  (adjusted  $R$ -) values are different from 0 for all treatments (Wilcoxon matched pairs rank test,  $p < 0.01$ ;  $N \geq 14$ ).

---

<sup>41</sup>Again, we do not distinguish between cases of rejected bribes and cases of defections.

<sup>42</sup>The variable  $R$  is a very crude measure of reciprocity in the sense that it does not control for potential determinants of the decision, e.g. the number of successful deals before  $t - 1$  or the magnitude of transfers in former decisions. Nevertheless it provides some (comparable) measure of  $B$ 's reciprocity.

<sup>43</sup>For groups without observations of both choices the interpretation of  $R$  is useless. Therefore we exclude those groups in the statistical tests.

Table 4: ***R*-value and adjusted *R*-value**

	IDT1	IDT2	TDT1	TDT2
<i>R</i> -Value	1.846*** (0.23)	2.036*** (0.36)	2.857*** (0.45 )	2.720*** (0.39)
adjusted <i>R</i> -Value	2.445*** (0.22)	2.300*** (0.32)	3.064*** (0.37)	3.142*** (0.28)

Standard deviations in parentheses

\*\*\* denotes significance of a Wilcoxon matched pairs rank sum test at the 1%-level.

We find substantially higher *R*-values for both GD treatments than for the ID treatments, which further strengthens the hypothesis concerning the success of strategies followed by groups of officials with respect to profit maximization through corrupt reciprocity (rent extraction). This also explains why corruption is more profitable for officials in the GD treatments (see Section 5.1.2). Type *B* subjects may anticipate type *O*'s strategies with respect to reciprocity and adapt by developing more accentuated reciprocity in their own behaviour.

We do not find any significant effect of gender (splitting observations according to *B*'s gender and applying pair-wise U-tests) with respect to bribers' reciprocity.

### 5.3 Switching behaviour and content analysis

So far we have analyzed revealed behaviour by considering the outcomes alone. Our data on the group decision making process allows for a more detailed analysis of the reasons and behavioural patterns underlying the treatment differences.

Since there are no significant differences in the behaviour of officials between the treatments in Stage 2 (see Section 5.1.2) we concentrate on *O*'s Stage 3 behaviour. The data on initial choices of individual officials in the GD treatments demonstrate that (at least part of) the higher levels of reciprocity within groups are due to mechanisms within the decision making process and not based on differences in individual preferences.

### 5.3.1 Switching behaviour

First, we identify situations in which officials within a group initially revealed opposing opinions on a decision, i.e. one official in the group chose to ‘defect’ and the other to ‘cooperate’ in the first step of the decision making process.

Second, we compare the (relative) numbers of successful corrupt transactions and failures following initial disagreement. For simplicity we pool cases of final disagreement and final consent against reciprocation (since both cases lead to a failed deal because of the veto power rule).

Tables 5 and 6 show average percentages of corrupt success and failure conditional on initial consent (or the lack of it).

Table 5: **Success and Initial Consent, TDT1**

	Successful corruption	Failed Corruption	Total
No initial Consent	23.13%	8.75%	31.88%
Initial Consent	23.75%	44.37%	68.12%
Total	46.88%	53.12%	100%

Averages are derived from 160 transactions (16 independent groups of officials in 10 periods) in TDT1

Table 6: **Success and Initial Consent, TDT2**

	Successful corruption	Failed Corruption	Total
No initial Consent	20.42%	18.33%	38.75%
Initial Consent	21.66%	39.59%	61.25%
Total	42.08%	57.92%	100%

Averages are derived from 240 transactions (12 independent groups of officials in 10 periods) in TDT2

Assuming independence of decisions (i.e. no influence of the process on final decisions) we would expect 100% of transactions without initial consent to fail because of the veto power of the non-reciprocating official. On the contrary, we find that the final decision was made in favour of (corruption-stabilizing) reciprocity in 72.6% (TDT1) and 52.70% (TDT2) of cases in which the two officials initially disagreed. Assuming that initial decisions reflect the true

underlying preferences<sup>44</sup>, this means that the decision making process alone is responsible for a large share of the treatment effects (higher success levels).

We conclude that (in both treatments) those officials who are in favour of engaging in, or maintaining, a successful corrupt relationship dominate the outcome of the decision making process although their decision-adversaries hold veto power. We take this finding as evidence for the persuasive argument theory (Pruitt 1971) which suggests that those participants (in the role of *O*) who provide the best ideas for maximizing long term individual payoffs during the experiment (which in our case is the maintenance of the corrupt relationship through reciprocity, see Section 5.2.) dominate the decisions within a group.

### 5.3.2 Content Analysis

In addition to the arguments derived from the comparison of outcomes between the treatments (see Section 5.1 and 5.2) and the analysis of choices in the different phases in the group decision making process (see Section 5.3), we are able to get some insight into the mechanism of group decision making by considering the content of the messages<sup>45</sup> exchanged during the group decision making processes of Stage 2 and Stage 3.

Only 22 (out of 28. 16 in TDT1 and 12 in TDT2) groups exchanged messages.<sup>46</sup> First we separate messages identifying 132 distinct statements.<sup>47</sup> We allocate each statement (sent in either of the two stages) into four main categories: ‘Neutral’ (Statements that do not contain any traceable argument, e.g. ‘Hello, nice game’); ‘Social’ (Arguments against the cooperation in the corrupt transaction mentioning the negative externality, e.g. ‘We have to consider the effect on the others, we should not cooperate’); ‘Strategic’ (arguments in favour of the stabilization of the reciprocal relationship with the objective of payoff maximization, e.g. ‘Let us cooperate, otherwise we won’t get any profit in the next period(s)’) and ‘Strategic Neg.’ (arguments against cooperation in a certain period to implicitly demand larger transfers in future periods, e.g. ‘Do not re-transfer, then he [the briber] will know to give more next time’).<sup>48</sup> We add a 5<sup>th</sup> category ‘Social/Strategic’ to account for (mostly twisted) statements that included both, other-regarding (social) and strategic (payoff maximizing) arguments.

---

<sup>44</sup>Although this is a very strong assumption it is not totally unreasonable to consider the hypothetical outcomes of successful corruption to be close to those observed in the ID treatments.

<sup>45</sup>We analyse all electronic chat messages exchanged by officials in the GD treatments.

<sup>46</sup>6 groups either did not encounter a situation of initial disagreement or ignored the possibility of writing messages

<sup>47</sup>A ‘conversation’ between two officials may yield more than one statement since it is split into single entries.

<sup>48</sup>All examples are translated (word by word) into English from the original statements in German.

Table 7 reports the relative frequencies of statements of the respective categories subdivided by the final outcome of the respective transaction in terms of success and failure.

Table 7: **Success and Content**

	Neutral	Social	Strategic	Strategic Neg.	Social/Strategic	Total
Success	8.3%	0.8%	27.3%	1.5%	6.8%	44.7%
Failure	31.8%	3.1%	12.1%	6.8%	1.5%	55.3%
Total	40.1%	3.9%	39.4%	8.3%	8.3%	100%

Percentages are derived from 132 distinct statements in TDT1 and TDT2

Of all statements, only 12.2% contain other-regarding arguments (Social and Social/Strategic). Their low frequency is noteworthy, and so is their lack of effectiveness (only 38% transaction finally fail). An explanation may be that in more than 75% of all situations a social argument was followed (in the same chat conversation) by a statement arguing in favour of strategic reciprocity. 82% of these situations ended with a successful corrupt transaction.

The majority (56%) of statements contains arguments in favour of some kind of strategic reciprocity.<sup>49</sup> In 19 (out of all 28 or 22 relevant) groups of officials we found at least one statement in favour of strategic reciprocity (positive or negative).

The dominance of arguments in favour of profit maximization is demonstrated not only by the relative frequency but also by the effectiveness as to corrupt success (83.4% of statements including an argument for strategic reciprocity ended in a successful corrupt transaction). This provides another piece of evidence for the hypothesis that the profit maximizing motive is the driving force in the decisions made in groups. Arguments that seem persuasive in the pursuit of profit maximizing are adopted and corresponding suggestions (i.e. maintenance of strategies aiming at payoff maximizing through corrupt reciprocity) realized, while arguments in favour of social efficiency (and over all fairness) are neglected, since they would lead to individually costly strategies. Again the argument is in line with the Persuasive Argument Theory (Pruitt 1982).

We leave it to further research to separate the effect of the decision making process from effects stemming exclusively from the nature of the exchange of arguments via electronic

---

<sup>49</sup>Additional to 63 statements of positive reciprocity we were able to identify 11 separate statements arguing in favour of strategic defection aimed at extracting larger bribes in future periods.

chat messages. For our purpose of evaluating the effectivity of the 4EP the effort of distinguishing between those two would lead to a more artificial setting and would not help to derive conclusions.

## 6 Conclusion

The results of our experiment are interesting in two respects. First, they serve as an assessment of the usefulness of the Four Eyes Principle. Second, they give some insight into the mechanism of group decision making.

With our experiment, using the framework of a simple 3-Stage game which is standard in the experimental corruption literature (Abbink et al 2002. Abbink 2004. Lambsdorff and Frank 2010), we show that the introduction of the Four Eyes Principle, which is generally promoted as one of the most effective tools to curb corruption (Poerting and Vahlenkamp 1998, Rieger 2005, Wiehen 2005), can be counterproductive. We find that it increases the relative number of successful corrupt transactions as well as the amount of bribes being transferred resulting in reduced welfare (measured by the sum of participants' payoffs) and equality (measured in the distribution of payoffs between participants).

Moreover we find two opposing effects of the introduction of the 4EP. One, the bribe splitting effect, is caused by the splitting of the transfer between two officials which reduces the level of corruption by changing the trade-off between its costs and its benefits. The other, the group decision making effect, increases the level of corruption. We are able to separate these two effects by the use of four different treatments and show that the group decision making effect over-compensates the bribe splitting effect. This leaves a negative total effect from the introduction of the 4EP with respect to the level of corruption and resulting social efficiency.

To explain the direction and magnitude of the group decision making effect (leading to higher conditional rates of reciprocity) we proceed in three steps. First, the differences in revealed strategies between groups and individuals (i.e. final outcomes) can be identified by non-parametric tests and quantified by (parametric) regression analyses of conditional levels of corruption. Groups of officials reciprocate more often for high transfers and less often for low transfers than individual officials. These 'sensible' (accentuated) strategies lead to a higher number of successful corrupt transactions in the group decision making treatments. Second, the analysis of behaviour within the group decision making process provides further

evidence. Contrary to predictions, in most cases initial disagreement between jointly deciding officials leads to a successful corrupt transaction despite the veto power of non-corrupt officials. Third we analyze the content of electronic chat-messages, exchanged during the decision making process. Arguments in favour of strategic reciprocity (i.e. initiating or maintaining only corrupt transaction that yield a large payoff in terms of high transfers) dominate the decision making process not only quantitatively but also in terms of effectiveness (outcomes).

The results of this 3-step analysis strongly suggest the dominance of the ‘profit maximizing motive’ in the group decision making process (Blinder and Morgan 2005, Kocher and Sutter 2005, 2007). This is in line with the persuasive argument theory (Pruitt 1971).

Moreover we show that the profit maximizing motive drives group decisions further away from the theoretical predictions in a situation where there is, in contrast to the gift exchange game (Fehr et al 1993), a strict trade-off between individual (unit-level) payoff maximization and social efficiency. While groups of officials seem to be better in maximizing their individual payoffs through maintaining a corrupt relationship based on trust and reciprocity (explainable by their higher cognitive capacity, Chalos and Pickard 1985), they do not manage to take the negative externalities (produced by successful corruption) into account, which ultimately aggravates the social dilemma of corruption.

Policies that prescribe group decision making should be restricted to situations for which the pursuit of maintaining reciprocal and payoff maximizing strategies are in line with the policies’ objectives. Our results cast serious doubt on the usefulness of the Four Eyes Principle for which this condition is not given.

We can show that by looking into the black box of the mechanism behind group decision making through the analysis of decision making processes in combination with content analysis within controlled laboratory experiments we may not only be able to help to interpret behavioural patterns but to derive characteristics describing situations for which the strategic use of group decision making could reduce social inefficiencies.



# Appendix

## Appendix A: Proofs

### Equilibrium in 3-Stage Game

Denote by  $I_{i,n}$  the information set in Stage  $i$  ( $i \in (1,3)$ ) of period  $n$  ( $n \in (1,10)$ ). Let  $p(I_{i,n})$  be the probability of reaching the respective stage and  $q('X'|I_{i,n})$  the conditional probability of the relevant agent choosing action 'X' once reached Stage  $i,n$ . An information set contains all relevant information about ego's and alter's behaviour up to the respective (relevant) stage. Furthermore let  $PO(I_{i,n})$  be the (sum of) payoff(s) gained up to the arrival of information set  $I_{i,n}$ .

First we show that there cannot be an equilibrium in which  $O$  chooses 'Cooperate' in Stage 3 of the last (10<sup>th</sup>) period.

Consider a Strategy-Set  $EQU1 = (s_{1,1}, s_{2,1}, s_{3,1}, s_{1,2}, \dots, s_{3,10})$  in which  $p(I_{3,10}) > 0$  and  $q('cooperate'|I_{3,10}) > 0$ . Compare the payoff, resulting from the realization of the Strategy-Set ( $PO(EQU1)$ ) to the one of an alternative set which consists of equal strategies up to  $I_{3,10}$  but for which  $q('cooperate'|I_{3,10}) = 0$  yielding payoff  $PO(EQU1_{new})$ . Since  $8+3*b < 12+3*b$ ,  $PO(EQU1) < PO(EQU1_{new})$  and hence  $EQU1$  cannot constitute an equilibrium (Stage-Game Nash Equilibrium).

Second we show that, in the last (10<sup>th</sup>) period,  $B$  will never choose ' $b > 0$ '.

Consider again a Strategy-Set  $EQU2 = (s_{1,1}, s_{2,1}, s_{3,1}, s_{1,2}, \dots, s_{3,10})$  in which  $p(I_{2,10}) > 0$ ,  $q('cooperate'|I_{3,10}) = 0$   $q('b > 0'|I_{3,10}) > 0$ . Again compare  $PO(EQU2)$  to  $PO(EQU2_{new})$ , the payoff of a strategy set that differs from the former only in  $q('b > 0'|I_{2,10}) = 0$ . Since  $12 - b \leq 12$ ,  $PO(EQU2) \leq PO(EQU2_{new})$  so that  $EQU2$  cannot constitute an equilibrium. Hence only a Strategy-Set featuring  $(s_{1,1}, \dots, s_{9,1}, 'b = 0', 'accept'/'reject', 'defect')$  can characterize an equilibrium.

Consider now a set  $S = \{k, \dots, 10\}$  of (the last 10-k) consecutive periods for which the above stated last period's equilibrium Strategy-Set is played. Assume  $q('cooperate'|I_{3,k-1}) > 0$  for the period  $k - 1$ . By the same line of arguments as for the last (10<sup>th</sup>) period we can easily repeat the task up to the point of excluding all Strategy-Sets that do not exhibit the strategy characteristics of the equilibrium in the 10<sup>th</sup> period. Letting  $k$  decrease from 10 down to 2. it is easy to see that the Stage-Game Nash Equilibrium remains the only Subgame-perfect Nash Equilibrium in the (finitely) repeated game.

## Responsibility and Veto Power

Consider two officials,  $O_i$  and  $O_j$ , who decide jointly in Stage 3 between ‘cooperate’ ( $c$ ) and ‘defect’ ( $d$ ). Consider  $O_i$ ’s preferences to be represented by the utility function  $U_i(a_1.a_2)$  where  $a_1$  is her own action and  $a_2$  that of  $O_j$ .  $O_i$ ’s decision is pivotal only if  $a_2 = c$ . If not, the outcome is  $(u,u)$  giving utility  $U_i(u,u)$ , independent from  $O_i$ ’s choice. Where in the ID treatments each official has intermediate responsibility (i.e. for the own deal only), the officials in the group decision making treatments have either no responsibility if paired with a non-reciprocal official or full responsibility if paired with a reciprocal one. This means that giving up responsibility works only in the cases where the socially optimal choice is taken anyway. Therefore, no official within a group can (for herself) deny responsibility for her group’s (corrupt reciprocal) behaviour since this would need the approval of both officials. For these reasons we believe that a lack of individual responsibility within a group cannot be applied as an argument in favour of the prediction of higher levels of corruptibility among officials within a group.

## Appendix B: Figures and tables

### Extensive forms of games in all treatments

In all treatments except TDT2 both,  $O$  and  $B$  decide once in every period. In TDT2 only  $B$  decides once per period while each  $O$  decides twice.

Figure 7: Extensive forms of TDT1 and IDT1

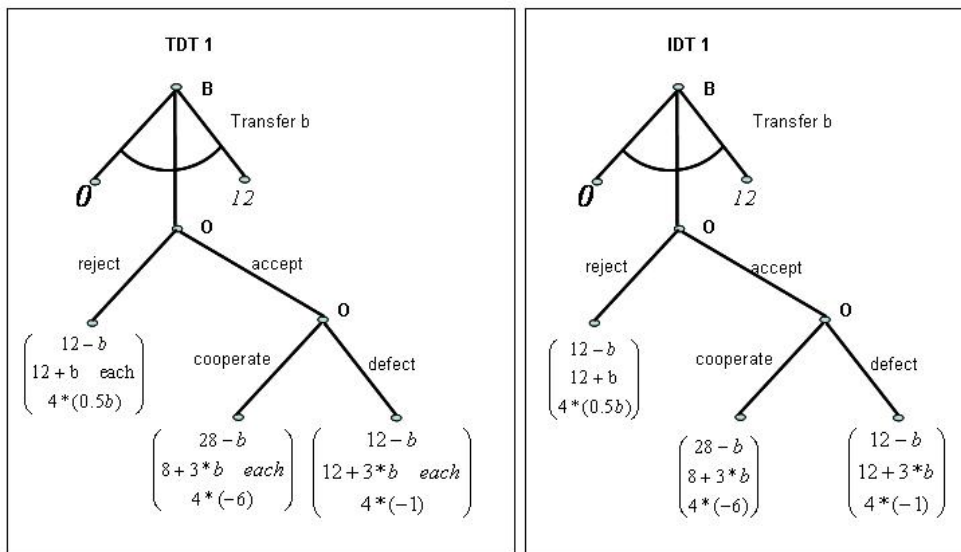
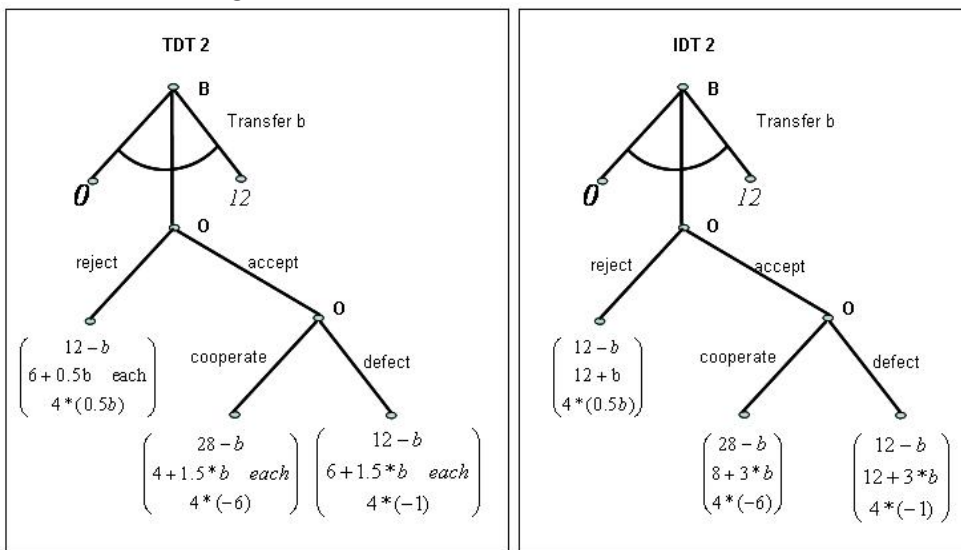


Figure 8: Extensive forms TDT2 and IDT2



Output table for Model M1 (Random Effects)

Table 8: **Model M1**

	Dependent variable: $PP$		
	Coefficients	Standard errors	
Constant	22.25***	1.02	Adjusted $R^2 = 0.54$ Number of subjects: 88 Number of periods: 10
$b$	1.89***	0.28	
$b^2$	-0.27	-0.29	
$D_{IDT2}$	-0.12	0.19	
$D_{TDT1}$	-6.34***	0.84	
$D_{TDT2}$	-3.07***	0.46	
$D_{IDT2} * b$	-0.13	0.17	
$D_{TDT1} * b$	2.05***	0.34	
$D_{TDT2} * b$	0.74***	0.21	
$D_{IDT2}^2 * b$	-0.02	0.14	
$D_{TDT1}^2 * b$	-0.25	0.21	
$D_{TDT2}^2 * b$	0.44	0.31	

Output table for Model M2 (OLS regression)

Table 9: **Model M2**

	Dependent variable: $Payoff$		
	Coefficients	Standard errors	
Constant	225.25***	54.92	$R^2 = 0.43$ Number of subjects: 80 Number of units: 64
$N$	15.89***	2.48	
$D_{IDT2}$	-7.12	10.29	
$D_{TDT1}$	-156.34***	42.87	
$D_{TDT2}$	-133.17***	39.72	
$D_{IDT2} * N$	-0.81	1.17	
$D_{TDT1} * N$	-3.14***	0.73	
$D_{TDT2} * N$	-5.03***	0.95	

\*\*\* denotes significance at the 1%-level, \*\* denotes significance at the 5%-level, and \* denotes significance at the 10%-level.

Output table for marginal effects in Probit Model M4

Table 10: Model M4

		Dependent variable: $SC$					
		Number of subjects/units: 96/64			Pseudo $R^2 = 0.33$		
Var	transfer level						
	0	1	2	3	4	5	
$b$	0.026***	0.034***	0.045***	0.055***	0.065***	0.073***	
$D_{IDT2} * b$	-0.002	-0.002	-0.003	-0.004	-0.004	-0.005	
$D_{TDT1} * b$	0.018***	0.025***	0.032***	0.039***	0.046***	0.052***	
$D_{TDT2} * b$	0.015***	0.020***	0.026***	0.032***	0.037***	0.043***	
$\hat{S}C_{IDT1}$	0.093	0.119	0.154	0.198	0.254	0.319	
$\hat{S}C_{IDT2}$	0.092	0.116	0.147	0.192	0.241	0.303	
$\hat{S}C_{TDT1}$	0.013	0.044	0.091	0.172	0.239	0.352	
$\hat{S}C_{TDT2}$	0.009	0.034	0.073	0.146	0.195	0.318	
Var	transfer level						
	6	7	8	9	10	11	
$b$	0.080***	0.083***	0.082***	0.078***	0.071***	0.062***	
$D_{IDT2} * b$	-0.005	-0.006	-0.006	-0.005	-0.005	-0.005	
$D_{TDT1} * b$	0.057***	0.059***	0.058***	0.056***	0.051***	0.044***	
$D_{TDT2} * b$	0.047***	0.048***	0.048***	0.045***	0.041***	0.036***	
$\hat{S}C_{IDT1}$	0.392	0.471	0.555	0.638	0.715	0.785	
$\hat{S}C_{IDT2}$	0.370	0.445	0.523	0.598	0.672	0.737	
$\hat{S}C_{TDT1}$	0.510	0.627	0.723	0.889	0.921	0.998	
$\hat{S}C_{TDT2}$	0.471	0.542	0.698	0.802	0.897	0.967	

\*\*\* denotes significance at the 1%-level, \*\* denotes significance at the 5%-level, and \* denotes significance at the 10%-level.

Marginal effects are calculated at the respective values of transfer and the means of the remaining independent variables.

## Appendix C: Instructions from TDT1 (translated into English)

### Instructions

In the next 90 minutes you will take part in an experiment in the laboratory of MELESSA. If you read the following instructions carefully, you can (depending on your decisions) earn money, additional to the show-up fee of 4 Euros. Additional to the money you can earn for yourself, you will affect the amount of donation to the public Aid organization ‘Doctors without Borders’. The money you will earn during the experiment will be added to the show-up fee and paid out in cash at the end of the experiment. The money that is going to be donated will be transferred to the donations account of ‘Doctors without Borders’.

During the experiment you are not allowed to communicate with the other participants. If you have questions, please approach one of the experimenters by raising your hand. In the case of violation of this rule we have to exclude you from any payments.

During the experiment we will refer to Experimental Monetary Units (EMU) instead of Euros. Your income will be calculated in EMU. In the end of the experiment the total amount will be exchanged in Euros.

The Exchange rate is  $1 \text{ EMU} = 5 \text{ Eurocents}$ .

All 24 participants are assigned to groups of four. Neither the experimenters nor the other participants know which group you are in. Your decisions stay completely anonymous.

### The Situation

There are two types in this experiment: type A and type B. The types play different roles and make decisions that affect their own income, the income of the other participants of the experiment and the amount of donation transferred to the organization ‘Doctors without Borders’. The type of a participant is allocated randomly.

A group of four consists of two type A and two type B participants who stay together for the entire experiment.

The experiment has 10 periods.

### Procedure:

All of the 10 periods consist of at most 3 Stages.

#### *Stage 1*

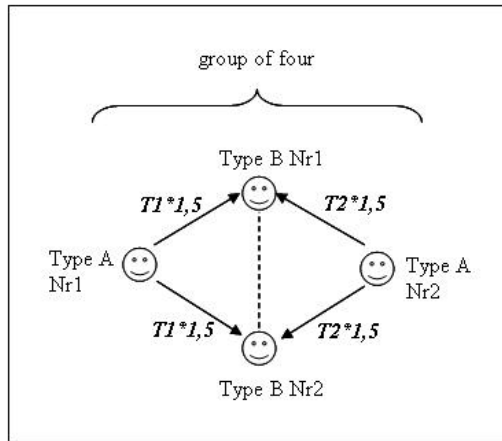
In the first Stage, every participant of type A (type A Nr 1 and type A Nr 2) decides on the size of their transfer (T1 denotes type A Nr 1’s transfer and T2 denotes type A Nr 2’s transfer) which has to lie between 0 and 12 EMU.

Next, the amount of the transfers is tripled and is split equally between the two type B participants (type B Nr 1 and type B Nr 2) of the group of four. If T1 is for example 6 EMU, type B Nr 1 receives 9 EMU ( $0.5 * 6 * 3 \text{ EMU}$ ) and type B Nr 2 9 EMU.

Hence there are 2 situations per group in any period:

Situation 1: Type A Nr1 transfers T1 to the two type B participants (where T1 is first tripled and then shared)

Situation 2: Type A Nr2 transfers T2 to the two type B participants (where T2 is first tripled and then shared)



### Stage 2

In Stage 2 the two type B participants decide jointly on how to react on the transfer of the respective type A participant. They have (in both situations) two alternatives.

1<sup>st</sup> Alternative: Both decide (for a specific transfer, e.g. T1) jointly for ‘keep’: In this case Stage 3 is entered

2<sup>nd</sup> Alternative: One or both decide in favour of ‘distribute’: In this case, the respective type A participant (e.g. type A Nr 1) does not get a bonus (and receives only  $12 - T1$  EMU). The type B participants both get 6 EMU plus half of the value of the transfer ( $6 + 0.5 * T1$  EMU). Moreover, the amount of  $2 * T1 + 24$  EMU is transferred as a donation to the organization ‘Doctors without Borders’.

A joint decision between the two subjects is found as follows.

First, each of the two type B participants decides individually whether to ‘keep’ or to ‘distribute’ the particular transfer.

If the decision is not unanimous (one type B participant wants to ‘keep’ and the other wants to ‘distribute’ the transfer), the decision of the fellow participant appears on his or her own screen.

Next, the participants decide once again separately. If there is still no agreement, the two type B participants can exchange messages via an electronic ‘chat’ (see explanation below) for one minute. After this the participants decide for the last time.

Note that only if both typeB participants decide in favour of ‘keep’ the third Stage is actually reached.

Since there are two type A participants in every group of four (type A Nr 1 and type A Nr 2), each of the type B participants has to decide (jointly with the other type B participants) in two situations: once for T1 and once for T2.

### Stage 3

In Stage 3 (which is only reached if both type B subjects have chosen ‘keep’) the two type B participants decide again jointly whether to initiate a re-transfer or not.

Again, both type B subjects decide separately first

If the decision is not unanimous (one type B participants wants to initiate the re-transfer and the other one does not), the decision of the other participant is shown on the screen. Then the participants can decide again separately. If there is still no consent, the participants enter again a ‘chat’ in which they can exchange

electronic messages for one minute.

After this, there is a final decision.

1. Case: Both type B participants decide in favour of a re-transfer. Both carry the costs of 2 EMU each (independent from the amount of the respective transfer). They both get 6 EMU plus one and a half times the value of the transfer less the costs of 2 ( $6 - 2 + 1.5 * T1$  EMU). The respective type A participant (type A Nr 1) receives a Bonus of 16 EMU in addition to the 12 EMU of initial endowment ( $16 + 12 - T1$ ). In this case there is no donation to the organization 'Doctors without Borders'.
2. Case: One or both type B participants decide against a re-transfer. In this case, there are no personal costs for the two type B participants (they get  $6 + 1.5 * T1$  each), the respective type A participant does not receive a bonus (and gets  $12 - T1$ ) and the donation to the organization is 20 EMU.

In the end, all participants are shown their personal income in the period. Please note, that the type A participants can thereby reconstruct whether or not the type B participants chose for or against the re-transfer. These (maximal) 3 stages are repeated 10 times (10 periods). Since the members of groups stay together, participants *always* interact with the same persons in the same roles for the entire experiment. (Type A Nr 1 remains type A Nr 1. type A Nr 2 remains type A Nr 2 etc.)

#### **Chat:**

Type B subjects potentially have the possibility to communicate via real time electronic messaging (Chat) with their fellow type B subject to agree on a joint decision (e.g. 'keep' or 'distribute') in Stage 2 and Stage 3.

The content of the communication is generally free to choose but there are some restrictions. You are not allowed to make statements about personal characteristics such as your name, age, address, gender, subject of study or any information that might lead to your identification. Moreover, strong language is strictly forbidden. Anyone who violates these rules of communication will be automatically expelled from the experiment and will not get any payments *for the entire experiment*.

Each participant in the chat can send as many messages to the other participant as he wishes or is able to send within the time limit of one minute.

Every message appears automatically on the screens of both type B participants of a group of four but cannot be seen by any other participant of the experiment.

#### **Payoff table**

The following table shows the kind of consequences the decisions of the participants lead to- in terms of their own payoff, the payoff of the other participants and the organization 'Doctors without Borders' (Example for T1).

The following table can be read as follows. Generally we start from the top and go down cell by cell. If a participant chooses a certain alternative, only those cells that lie directly beneath it are relevant for the next period.

*The payment table is analogous for situations in which T2 (Transfer of Type A Nr2) is relevant.*

Note that each type B participant receives two payments because two situations are relevant for each of



1. Stage	Type A Nr1 chooses <b><i>T1</i></b>		
2. Stage	At least one Type B participant (Type B Nr1 <b>or</b> Type B Nr2 <b>or both</b> ) decide in favour of ' <i>distribute</i> ' for <b><i>T1</i></b>	<b>Both</b> Type B participants (Type B Nr1 <b>and</b> Type B Nr2) decide in favour of ' <i>keep</i> ' for <b><i>T1</i></b>	
3. Stage	Stage 3 is not reached	Type B Nr1 or Type B Nr2 or both decide against a <b>re-transfer</b> to Type A Nr1	<b>Both</b> (Type B Nr1 and Type B Nr2) decide in favour of a <b>re-transfer</b> to Type A Nr1
Payoff Type A Nr1	<b>12 - T1</b>	<b>12 - T1</b>	$12 + 16 - T1 =$ <b>28 - T1</b>
Payoff Type B Nr1/Type B Nr2	<b>6 + 0,5*T1</b>	<b>6 + 1,5*T1</b>	$6 - 2 + 1,5*T1 =$ <b>4 + 1,5*T1</b>
<b>Donation to 'Doctors without Borders'</b>	<b>2*T1+ 24</b>	<b>20</b>	<b>0</b>

them, one with type A Nr 1 (T1 is relevant) and one with type A Nr 2 (T2 is relevant). These two are added up for any period.

For the type A participants only one situation per period is relevant so that there is only one payment per period.

Each participant gets information at the end of each period about his own personal payoff. Type A subject can infer whether type B subjects have chosen to initiate a re-transfer or not.

Note that the sum of payments (exchanged in Euros) to the organization is actually donated to 'Doctors without Borders'.

### Timing

**Stage 1:** Type A Nr1 chooses T1 and Typ A Nr2 chooses T2

**Stage 2:** Type B Nr1 and Type B Nr2 decide (T1 and T2) each time jointly about 'keep' or 'distribute'

**Stage 3:**

Situation (for T1): Only reached if in this situation, both Type B participants chose 'keep' in Stage 2. In this case, both Type B participants decide jointly whether to initiate a re-transfer or not.

Situation (for T2): is analogous, only for T2.

In the end of each of the ten periods, each participant gets information about his/her own payoff in the respective period.

At the end of the last (10<sup>th</sup>) period, participants get to know their final income and their payment in Euros.

The following control questions will help you to get a better understanding of the situation. All the necessary information can be found in the payment table.

Please answer all the control questions and raise your hand when you have finished. An experimenter will come to your place to check your solutions.

### Question 1

Assume you are type A Nr1 and you chose a transfer of 4 EMU (T1). The other participant of type A (type A Nr2) has chosen a transfer (T2) of 10 EMU.

Situation 1 (T1): One of the participants of type B in your group of four (type B Nr1) decides to 'distribute' your transfer (T1). The other participant of type B (type B Nr2) wants to 'keep' your transfer (T1). (Therefore Stage 3 is not reached.)

Situation 2 (T2): Both type B participants chose to 'keep' the transfer of type B Nr 2 (T2) in Stage 2 and decide against a re-transfer in Stage 3.

a) What is the payoff of type B Nr1 in the situation (1) with you?

Your answer:.....

b) What is the payoff of type B Nr2 in the situation (1) with you?

Your answer:.....

c) What is your (type A Nr1) total payoff in this period?

Your answer:.....

d) What is the total payoff of type B Nr1 for all situations relevant to him/her?

Your answer:.....

e) What is the total payoff of type B Nr2 for all situations relevant to him/her?

Your answer:.....

f) What is the total payoff of type B Nr2 in this period?

Your answer:.....

g) What is the amount of donation to 'Doctors without Borders' caused by the situation relevant to you (type A Nr1)?

Your answer:.....

h) What is the amount of donation to 'Doctors without Borders' caused by the situation relevant to type A Nr2?

Your answer:.....

i) What is the total amount of donation to 'Doctors without Borders' in this period?

Your answer:.....

j) What is the total amount of payoff generated by the decisions of your group of four?

Your answer:.....

## Question 2

Assume that you (type A Nr1) and type A Nr2 have both chosen a transfer of 0 (T1 is 0 EMU and T2

is 0 EMU). Neither participant of type B (neither type B Nr1 nor type B Nr2) wants to 'keep' any of the two transfers in Stage 2.

a) What is your (type A Nr1) total payoff in this period?

Your answer:-----

b) What is the total payoff of type A Nr2 in this period?

Your answer:-----

c) What is the total payoff of type B Nr1 in this period?

Your answer:-----

d) What is the total payoff of type B Nr2 for all situations relevant to him/her?

Your answer:-----

e) What is the total amount of donation to 'Doctors without Borders' in this period?

Your answer:-----

f) What is the total amount of payoff generated by the decisions of your group of four?

Your answer:-----

3. Question Assume that you (type A Nr1) have chosen a transfer of 5 EMU (T1 is 5 EMU) and type A Nr2 has also chosen a transfer of 5 EMU (T2 is 5 EMU). Both participants of type B (type B Nr1 and type B Nr2) decide to 'keep' the transfer and initiate a re-transfer in Stage 3.

a) What is your (type A Nr1) total payoff in this period?

Your answer:-----

b) What is the total payoff of type A Nr2 in this period?

Your answer:-----

c) What is the total payoff of type B Nr1 in this period?

Your answer:-----

d) What is the total payoff of type B Nr2 for all situations relevant to him/her?

Your answer:-----

e) What is the total amount of donation to 'Doctors without Borders' in this period?

Your answer:-----

f) What is the total amount of payoff generated by the decisions of your group of four?

Your answer:-----

## References

- Abbink, K., B. Irlenbusch, and E. Renner, E. (2000), “The Moonlighting Game”, *Journal of Economic Behavior and Organization* VOL 22. pp 265–77
- Abbink, Klaus, Bernd Irlenbusch, and Elke Renner (2002), An Experimental Bribery Game, *The Journal of Law, Economics and Organization* VOL 18/2. pp 428–54
- Abbink, K. (2004), “Staff rotation as an anti-corruption policy: an experimental study”, *European Journal of Political Economy*, VOL 20. pp 887–906
- Abbink, K. and H. Hennig-Schmidt (2006), “Neutral versus loaded instructions in a bribery experiment”, *Experimental Economics* VOL 9/2. pp 103–21
- Abbink, K. (2006), “Laboratory experiments on corruption”, in: *International Handbook on the Economics of Corruption*, edited by S. Rose-Ackerman, Cheltenham, pp. 418–37
- Aidt, T.S. (2003), “Economic Analysis of Corruption: A Survey”, *The Economic Journal*, VOL 113. pp 632–52
- Alatas, V., L. Cameron, A. Chaudhuri, N. Erkal, L. Gangadharan (2006), “Gender and Corruption: Insights from an Experimental Analysis” Department of Economics - Working Papers Series 975. The University of Melbourne
- Andreoni, J., J. Miller (2002), “Giving according to GARP: An experimental test of the consistency of preferences for altruism”, *Econometrica* VOL 70. pp 737–53
- Armantier, O. A. Boly (2008), “Can Corruption Be Studied in the Lab? Comparing a Field and a Lab Experiment”, Cirano Working paper Series, Quebec
- Bardhan, P. (1997), “Corruption and Development: A Review of Issues”, *Journal of Economic Literature*, VOL 35/3. pp 1320–46
- Becker, G.S. (1968), “Crime and punishment: an economic approach”; *Journal of Political Economy*, VOL 78. pp 169–217
- Bertrand, M., S. Djankov, R. Hanna, S. Mullainathan (2007), “Obtaining a Driver’s License in India: An Experimental Approach to Studying Corruption” *The Quarterly Journal of Economics*, VOL 122/4. pp 1639–76
- Bishop G., D. Myers (1974), “Informational influences in group discussion”, *Organizational Behavior and Human Performance* VOL 12. pp 92–104

- Blinder, A., J. Morgan (2005), “Are two heads better than one? An experimental analysis of group vs. individual decision making”, *Journal of Money Credit and Banking*, VOL 37. pp 789–812
- Bornstein G., I. Yaniv (1998), “Individual and group behavior the ultimatum game: are groups more rational players?”, *Experimental Economics* VOL 1. pp 101–08
- Buechner S, A. Freytag, L. G. Gonzalez, W. Gth (2008), “Bribery and public procurement: an experimental study”, *Public Choice* VOL 137. pp 103–11
- Burnstein E., A. Vinokur, Y. Trope (1973), “Interpersonal comparison versus persuasive argument: a more direct test of alternative explanations for group-induced shifts in individual choices”, *Journal of Social Psychology* VOL 9. pp 236–45
- Cameron, L. A. Chaudhuri, N. Erkal, L. Gangadharan (2006), “Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students”, Working Papers Series 975. The University of Melbourne
- Cameron, L. A. Chaudhuri, N. Erkal, L. Gangadharan (2005), “Do attitudes towards corruption differ across cultures? Experimental evidence from Australia, India, Indonesia and Singapore”, Working Papers Series 975. The University of Melbourne
- Cameron, L. (2004), “Corruption: An experimental Analysis”, Working Papers Series 975. The University of Melbourne
- Cason, T., V. Mui (1997), “A laboratory study of group polarization in the team dictator game”, *Economic Journal* VOL 107. pp 1465–83
- Chalos P., S. Pickard (1985), “Information choice and cue use: An experiment in group information processing”, *Journal of Applied Psychology* VOL 70. pp 634–41
- Charness, G., M. Rabin (2002), “Understanding social preferences with simple tests”, *Quarterly Journal of Economics* VOL 117. pp 817–69.
- Cox, J.C. Friedman, D. Friedman, S. Gjerstad, (2007), “A tractable model of reciprocity and fairness”, *Games and Economic Behavior*, VOL 59/1. pp 17–45
- della Porta, D., A. Vanucci (1999), “Corrupt Exchanges, Actors, Resources and Mechanisms of Political Corruption”, de Gruyter, New York

- Dollar, D., R. Fisman, R. Gatti, (2001), “Are women really the fairer sex? Corruption and women in government”, *Journal of Economic Behavior and Organization* VOL 46. pp 423–29
- Dusek, L. A. Ortmann, L. Lizal (2005), “Understanding Corruption and Corruptibility through Experiments”, *Prague Economic Papers* VOL 2. pp 147–62
- Falk, A., U. Fischbacher (2006), “A theory of reciprocity”, *Games and Economic Behavior* VOL 54/2. pp 293–315
- Fehr, E., U. Fischbacher (2003), “The nature of human altruism”, *Nature* VOL 425. pp 785–91
- Fehr, E., S. Gächter (1998) “Reciprocity and economics: The economic implications of Homo Reciprocans”, *European Economic Review* VOL 42. pp 845–59
- Fehr, E., S. Gächter, G. Kirchsteiger (1997), “Reciprocity as a contract enforcement device: experimental evidence”, *Econometrica* VOL 65/4. pp 833–60
- Fehr, E., G. Kirchsteiger, A. Riedl (1993), “Does fairness prevent market clearing? an experimental investigation”, *Quarterly Journal of Economics* VOL 108. pp 437–59
- Fehr, E., K. Schmidt (1999), “Theory of Fairness, Competition and Cooperation”, *Quarterly Journal of Economics* VOL 114. pp 817–68
- Fehr, E., K. Schmidt (2006), “The Economics of fairness, reciprocity and altruism - experimental evidence and new theories”, *Handbook on the Economics of Giving, reciprocity and Altruism* VOL 1. pp 615–91
- Fischbacher, U. (2007), “Z-Tree: Zurich Toolbox for Ready-made Economic Experiments”, *Experimental Economics* VOL 10/2. pp 171–78
- Frank, B. and G. Schulze (2000), “Does Economics Make Citizens Corrupt?”, *Journal of Economic Behavior and Organization*, VOL 43/1. pp 101–13
- Fudenberg, D., J. Tirole (1991), “Game Theory”, The MIT Press, Cambridge Massachusetts/London England
- Fukuyama, F. (1995), “Trust: the social virtues and the creation of prosperity”, Free Press, New York
- Gibbons, R. (1992), “A Primer in Game Theory”, Harvester, Wheatsheaf

- Gneezy, U. and J. A. List (2006), “Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments”, *Econometrica*, VOL 74. pp 1365–84
- Groenendijk, N. (2004), “A principal-agent model of corruption”, *Crime, Law and Social Change* VOL 27/3.4. pp 207–29
- Heckman, J. (1979), “Sample Selection Bias as a Specification Error”, *Econometrica*, VOL 47. pp 153–61
- Huntington, S., (1968), “Political order in changing societies”, Yale University Press, New Haven
- Hussein, A. (2005), “How the new law on public procurement implements the anti-corruption recommendations of the Supreme Chamber of control”, *Fighting Corruption and Promoting Integrity in Public Procurement*, OECD Publishing, Paris
- Jacquemet N. (2005), “Corruption as Betrayal: Experimental Evidence“, GATE Working Paper nr 2005-06
- Klitgaard, R. (1988), “Controlling corruption”, University of California Press
- Knapp, W. (1986), “Sittenwidrige Vermittlung von Regierungsauftraegen durch Bestechungsgelder”, *Recht der Internationalen Wirtschaft*, RIW, pp 999–1001
- Kocher M., M. Sutter (2007), “Individual versus group behavior and the role of the decision making procedure in gift-exchange experiments”, *Empirica* VOL 34. pp 63–88
- Kocher M., S. Strau, M. Sutter (2006), “Individual or team decision making Causes and consequences of self-selection”, *Games and Economic Behavior* VOL 56. pp 259–70
- Kocher M., M. Sutter (2005), “The decision maker matters. Individual versus team behavior in experimental beauty-contest games”, *Economic Journal* VOL 115. pp 200–23
- Kugler, T., G. Bornstein, M. Kocher, M. Sutter (2007), “Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy”, *Journal of Economic Psychology* VOL 28. pp 646–57
- Lambert-Mogiliansky, A., M. Majumdar, R. Radner (2006), “Petty corruption: a game-theoretic approach”, *International Journal of Economic Theory* VOL 4/2. pp 273–97
- Lambsdorff, J. Graf (2002), “Corruption and Rent-Seeking”, *Public Choice*, VOL 113/1.2. pp 97–125

- Lambsdorff, J. Graf, M. Taube, M. Schramm (2005), “The New Institutional Economics of Corruption”, Routledge, London and New York
- Lambsdorff, J. Graf (2007), “The Institutional Economics of Corruption and Reform, Theory, Evidence and Policy”, Cambridge University Press
- Lambsdorff, J. Graf, M. Nell (2007), “Fighting Corruption with Asymmetric Penalties and Leniency”, CEGE Discussion Paper ISSN 1439–2305
- Lambsdorff, J. Graf, B. Frank (2007), “Corrupt Reciprocity - an Experiment”, Diskussionsbeitrag Nr. V-51-07. Volkswirtschaftliche Reihe ISSN 1435-3520
- Lambsdorff, J. Graf, B. Frank (2010), “Bribing versus gift-giving An experiment”, Journal of Economic Psychology, VOL 31/3. pp 347–57
- Lambsdorff, J. Graf, B. Frank (forthcoming), “Gender Effects in Laboratory Corruption Experiments”, Forthcoming in European Journal of Development Research
- Leff, N. (1964), “Economic development through bureaucratic corruption”, American Behavioral Scientist, VOL 8/3. pp 8–14
- Levine, D. (1998), “Modeling altruism and spitefulness in experiments”, Review of Economic Dynamics, VOL 1. pp 593–622
- Levine, J., R. Moreland (1998), “Small groups” in ‘Handbook of social psychology’, VOL, pp 415–69
- Mauro, P. (1995), “Corruption and Growth”, Quarterly Journal of Economics, VOL 110. pp 681–712
- Olken, B. A. (2007), “Monitoring Corruption: Evidence from a Field Experiment in Indonesia”, Journal of Political Economy VOL 115/2. pp 200–49
- OECD (2010), “Guidelines for Managing Conflict of Interest in the Public Service”, <http://www.oecd.org/dataoecd/31/15/36587312.pdf>
- Pechlivanos, L. (1997), “Self-Enforcing corruption and Optimal Deterrence”, GREMAQ Working Paper Series 97.18.460
- Pereira, P.T., N. Silva, J. Andrade Se ilva (2006), “Positive and negative reciprocity in the labor market”, Journal of Economic Behavior & Organization, VOL 59. pp 406–22



- Poerting, P., W. Vahlenkamp (1998), “Internal strategies against corruption: guidelines for preventing and combatting corruption in police authorities”, *Crime Law and Social Change*, VOL 29. pp 225–49
- Pruitt, D.G. (1971), “Choice Shifts in Group Discussion: An Introductory Review”, *Journal of Personality and Social Psychology*, VOL 20/3. pp 339–60
- Rieger H.-J. (2005), “Prevention a key factor in fighting corruption”, United Nations Public Administration Network
- Rivas, M.F. (2007), “An experiment on corruption and gender”, revised version, October 2008. mimeo
- v. Rose-Ackerman, S. (1999), “Corruption and Government, Causes, Consequences and Reform” in ‘The New Institutional Economics of Corruption’, Cambridge University Press 1999
- v. Rose-Ackerman, S. (2006), “International handbook on the economics of corruption2. Marston Book Services Limited, Abingdon, Oxon OX14 4YN UK
- Roth A., I. Erev (1995), “Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term”, *Games and Economic Behaviour* VOL 8. pp 164–212
- Schubert, R. M. Brown, M. Gysler, H.W. Brachinger (2009), “Financial Decision-Making: Are Women really more risk-averse?”, *The American Economic Review*, VOL 89/2. pp 381–5
- Schulze, G. and B. Frank (2003), “Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility”, *Economics of Governance* VOL 4/2. pp 143–60
- Sung, H.E. (2003), “Fairer Sex or Fairer System? Gender and Corruption Revisited”, *Social Forces* VOL 82. pp 703–23
- Swamy, A., Knack, S., Lee, Y., and Azfar, O. (2001), “Gender and Corruption”, *Journal of Development Economics* VOL 64. pp 25–55
- Tajfel, H., J.C. Turner (1986), “The social identity theory of intergroup behavior”, in S. Worchel and W.G. Austin: *Psychology of intergroup relations*, pp 7–24
- Tonoyan, V. (2005) “The Dark Side of Trust: Corruption and Entrepreneurship”, Hhmann, H. and Welter, F. (eds): ‘Trust and Entrepreneurship: A West-East-Perspective’, Cheltenham, Edward Elgar, pp 39–58

Wiehen, M. (2005), "Global Corruption Report 2005", Transparency International, Pluto Press, London and Ann Harbor